Experimental programme

Practical skills · SL and HL

Counting birds since 1900

> Careful, repeatable methods turn observations into evidence.

- Audubon reports 27 participants and 25 counts in 1900.
- Effort, detectability, weather and coverage affect raw totals.
- Numerical practical examples here are invented unless sourced.

A focused research question

Variable	**Lichen example**
Independent	Distance from the road
Dependent	Lichen species per trunk
Controlled	Tree species, trunk height, aspect, date

Hypothesis and null hypothesis

Hypothesis

A testable prediction with a reason.

Null hypothesis

A specified no-association or no-difference model.

Safety and ethics

> Plan for safety, the environment and the people you study before you start.

- Risk assessment; minimise trampling and disturbance.
- **Informed consent**, anonymity, the right to withdraw.

What is the independent unit?

Repeated measurements

Ten readings on one tree remain one sampled tree.

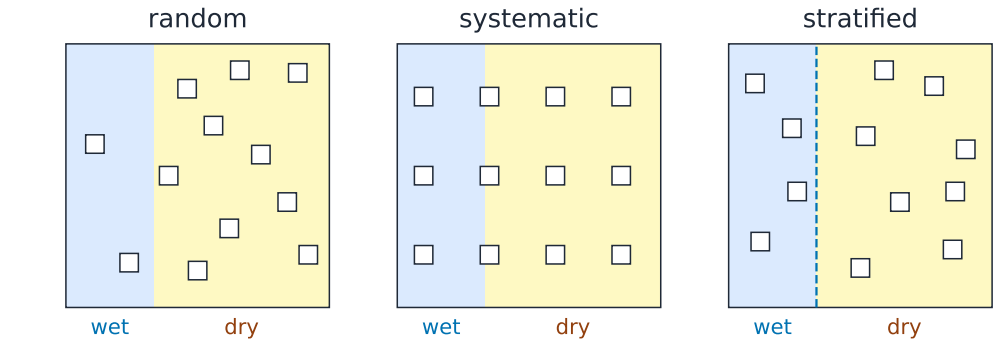
Repeated samples within one stream do not create ten independent streams.

Independent replication

Choose units at the level of the **research question**.

Account for clustering and spatial dependence.

Sampling strategies



Match selection and weighting to the target population.

Scaling up from quadrats

1. **Given.** Ten 1 m^2 quadrats contain 50 orchids in total; meadow area 2500 m^2.
2. **Mean density.** $50 \div 10 = 5$ plants per m^2
3. **Population.** $5 \times 2500 = 12\,500$ plants
4. **Assumption.** The quadrats represent the whole meadow

Decision

Estimates are only as good as the **sample**

Measuring abiotic factors

Factor	Method
Dissolved oxygen	Calibrated probe or suitable titration
Loss on ignition	Mass lost as fraction of initial dry soil
Transparency	Secchi depth; not directly NTU
Air pollution	Direct measurements plus contextual biological evidence

Social science methods

Closed questions

Yes or no, rating scales.

Easy to analyse.

Open questions

Richer answers.

Harder to compare.

Choose sampling effort using a pilot

Different curves

Species-area: increase the sampled area.

Accumulation: add samples or effort.

Limits of a plateau

A plateau is conditional on habitat, season and method.

Use variability and required precision as well as practical limits.

Weight the strata

1. **Wet third** Density 2 plants/m^2; contribution $= 2/3$ plants/m^2.
2. **Dry two-thirds** Density 8 plants/m^2; contribution $= 16/3$ plants/m^2.
3. **Whole field** Add contributions: 6 plants/m^2, not the unweighted mean 5.

Decision

The target area proportions determine the weights.

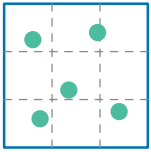
Soil mass losses need different denominators

1. **Net masses** Wet 30 g; dry 24 g; ignition residue 21.6 g.
2. **Water** Wet-basis $6/30 = 20\%$; dry-basis $6/24 = 25\%$.
3. **Ignition** $2.4/24 = 10\%$ of dry mass, a method-dependent proxy.

Decision

Loss on ignition is not a direct measurement of organic carbon.

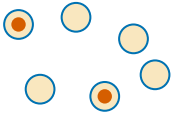
Three field sampling methods



quadrat
count species in a square



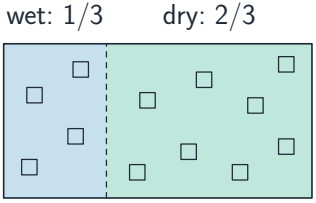
transect
record along a line



mark-recapture
mark some (red), release, then re-catch

Choose the method that fits the organism

Weighting strata



weighted: $\frac{1}{3}(2) + \frac{2}{3}(8) = 6$ m^{-2} ✓

unweighted pooling of equal samples per habitat: 5 m^{-2} ×

(invented meadow)

Pooling equal samples would give 5

Random or systematic error?

Random

Readings scatter both ways.

A fluctuating probe.

Systematic

Every reading shifted the same way.

An uncalibrated pH probe.

Label a half-range convention

1. **Given.** Invented dissolved-oxygen readings: 8.2, 8.6, 8.0, 8.4, 8.3 mg/L.
2. **Mean.** $41.5 \div 5 = 8.3$ mg/L
3. **Half-range.** $(8.6 - 8.0) \div 2 = 0.3$ mg/L
4. **Percentage.** $0.3 \div 8.3 \times 100 \approx 3.6\ \%$

Decision

8.3 ± 0.3 mg/L, explicitly mean $\pm$ half-range

Reliability and validity

Reliability

Would repeats agree?

More samples, consistent method.

Validity

Does it measure the right thing?

Does it test the question fairly?

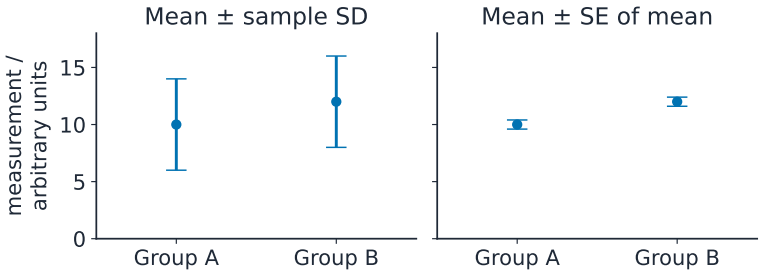
SD and SE answer different questions

1. **Squared deviations** Five oxygen readings: sum $= 0.20$ (mg/L)^2.
2. **Sample SD** $\sqrt{(0.20/4)} \approx 0.224$ mg/L describes observed spread.
3. **SE of mean** For independent observations, SD/$\sqrt{5} = 0.100$ mg/L.

Decision

Half-range 0.3, SD 0.224 and SE 0.100 are distinct quantities.

Read the error-bar definition



Invented summaries: n = 100 per group; SD = 4
SE = 4 / √100 = 0.4. Identify bars before interpreting.

Overlapping SD bars are not a universal significance test.

Propagate the stated kind of uncertainty

Worst-case bounds

Add absolute bounds for a sum or difference.

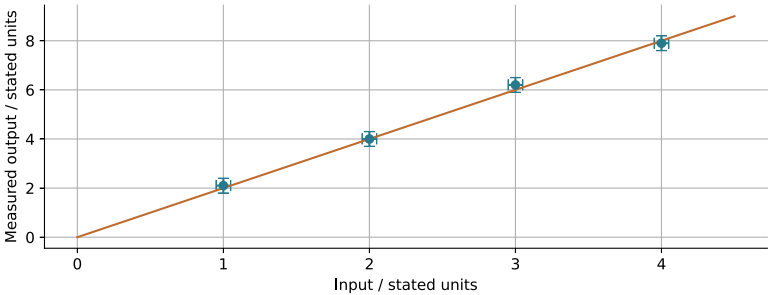
For small product/quotient bounds, add fractional bounds.

Random standard uncertainties

Independent random standard uncertainties often combine in quadrature.

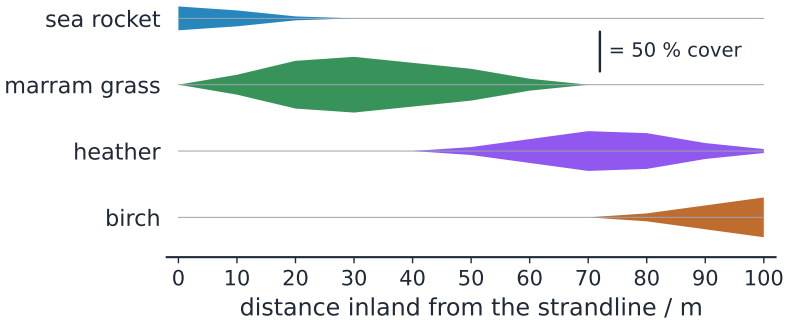
Correlations can change the rule; state the assumptions.

Uncertainty bars and a trend



The data support proportionality within their **uncertainty**

Kite diagrams



Distance is a spatial axis, not a measured time sequence.

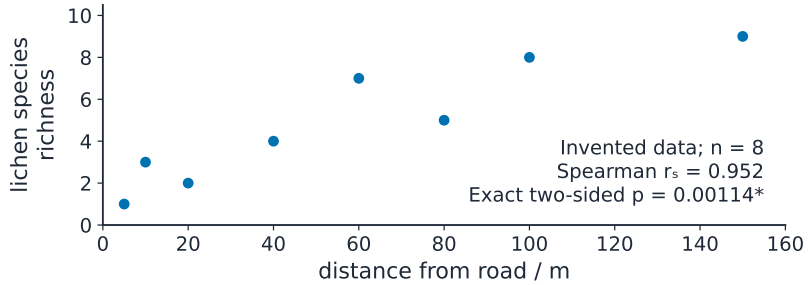
Spearman's rank

- Untied lichen ranks: $\Sigma D^2 = 4$, $n = 8$.
- $r_s = 20/21 \approx 0.952$; critical $|r_s| = 0.738$ at two-sided $\alpha = 0.05$.

$$r_s = 1 - \frac{6 \sum D^2}{n(n^2 - 1)}$$

Conditional inference: reject the specified null if design assumptions hold.

A coefficient is not a p-value



*Independent, exchangeable ranks under the null.
The p-value is not the probability that the null is true.

A small p-value does not identify a cause or repair the sampling.

Check: what can you conclude?

Your turn

Spearman's coefficient is 0.95 for lichens against distance from a road. Which conclusion is justified?

A. The data alone prove traffic pollution kills lichens
B. A positive rank association, conditional on stated test assumptions
C. Distance itself causes every lichen to grow
D. The null has probability 0.05 of being true

Alpha is a threshold; p comes from the data

Decision rule

Choose α and one- or two-sided testing before evaluating results.

A critical value is not the measured p-value.

Conditional probability

p concerns statistics at least as extreme under the specified null.

It is not P(null true), effect size or ecological importance.

A negative rank association

1. **Invented trampling data** Six pairs give sum of squared rank differences 68.
2. **Coefficient** $1 - 408/210 = -33/35 \approx -0.943$.
3. **Two-sided decision** $|r_s|$ exceeds 0.886 at $\alpha = 0.05$ under the stated model.

Decision

Keep the negative direction; association is not a causal demonstration.

Evaluating the method

For each weakness, give its effect on the results and a realistic improvement.

- Mixed tree species can confound the distance comparison.
- **Sample** comparable independent trees; record relevant covariates.

Applying the results

> Keep claims in proportion to the evidence.

- Link evidence, limits and stakeholder values to the recommendation.
- AHL: law, economics and ethics lenses can inform the application.

Report a null result honestly

What the test says

Failing to reject means insufficient evidence against this null in this test.

Report estimates, **uncertainty**, **sample** size and assumptions.

What it does not say

It does not prove no relationship or no ecological importance.

Nonlinear patterns, low power and poor measurement can limit inference.

Exam tips

> A focused **research question** names both variables and the place.

- Justify sampling strategy and size.
- Compare the coefficient with the critical value.
- Never claim causation from **correlation** alone.

Check yourself

Try these

Ten 1 m^2 quadrats in a 4000 m^2 field hold 30 dandelions. Estimate the population.

Distinguish between random and systematic errors.

Explain why correlation does not prove causation.