Lecture 1

Basic Principles of Experimental Designs

Norberto E. Milla, Jr.

2025-08-13

What is an experiment?

- an investigation set up to provide answers to question(s) of interest:
  - Is any reduction in disease infection achieved with a new 'resistant' variety compared with a standard 'control' variety?
  - How do plant metabolites respond to increasing drought stress at different stages of development?
  - Which chemicals, of several under study, show insecticidal activity?
  - How is yield related to plant spacing, and does this relationship vary between varieties?
- involves *manipulation* of the levels or settings of the independent variable(s)
- usually involves comparison of treatments
- it is useful to establish cause-and-effect

Unlocking of terms

Response - the dependent variable; observation or measurement to be taken

- Survival rate
- species diversity
- dissolved amount of a compound (Chemistry)
- corn yield (Agriculture)
- quality of chicken meat (Animal Science)
- compressive strength of concrete (Civil Engineering)
- score of a student in a test (Education)

Unlocking of terms

Factor - an independent variable whose effect on a measured response we want to study; it refers to the experimental conditions to be tested

- varieties
- nitrogen rates
- chemical compounds

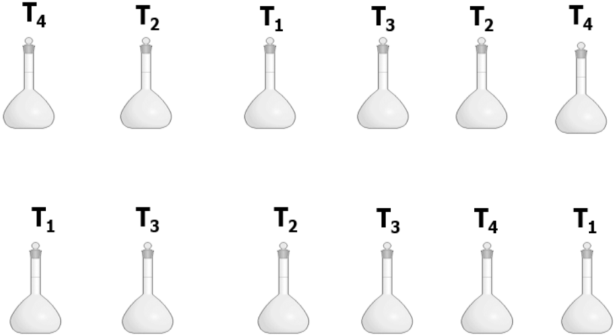
Treatment - different settings of a factor, or combinations of the levels of 2 or more factors

- varieties (V): V_1, V_2, V_3
- nitrogen rates (N): N_1, N_2
- Variety x Nitrogen rate combination: V_1N_1, V_2N_1, V_3N_1, V_1N_2, V_2N_2, V_3N_2

Unlocking of terms

Experimental unit - the unit of experimental material to which a treatment is applied

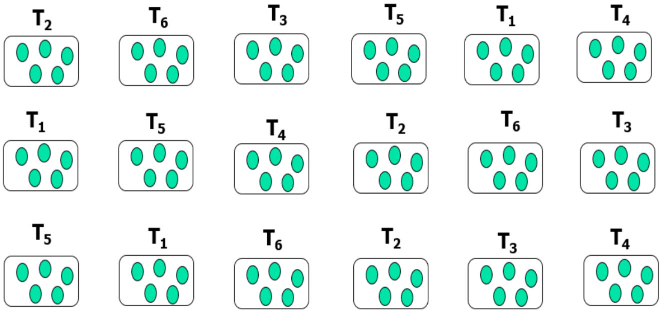
- Experiment: Effect of various temperature levels (T_1, T_2, T_3, T_4) on the yield of a chemical process



Unlocking of terms

Experimental unit - the unit of experimental material to which a treatment is applied

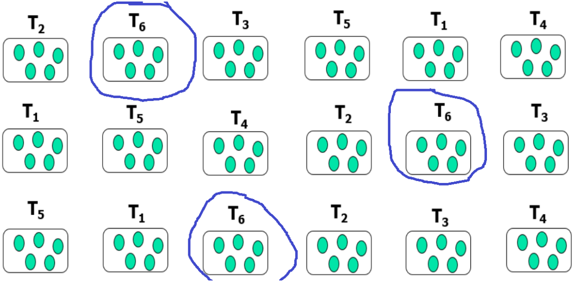
- Experiment: Effect of type of feed and amount of water (T_1, T_2, T_3, T_4, T_5, T_6) on the quality of chicken meat



Unlocking of terms

Measurement unit - physical entity on which a measurement is taken; may not be the same as the experimental unit - Individual flask - Sample of meat of individual chicken

Experimental error - observed differences in the responses of experimental units given the same treatment; a measure of precision of the experiment



Unlocking of terms

Possible causes of experimental error

- the natural differences in the experimental units prior to their receiving the treatment
- the variation in the devices that record the measurements
- the variation in setting the treatment conditions,
- the effect on the response variable of all extraneous factors other than the treatment factors

Unlocking of terms

Experimental layout - the placement of experimental treatments on the experimental units whether it be over space, time, or type of material

Linear model - mathematical representation of the relationship of the response variable with the factor(s) in the experiment and the experimental error

Types of factors:

- *Fixed*: the levels of the factor are selected on purpose or its levels consist of all possible levels of that factor
- *Random*: the levels of the factor used in the experiment consist of a random sample of potential levels within a well defined population

Types of linear models

Fixed model

- all factors in the experiment are fixed
- inference in restricted to factor levels actually used in the experiment
- parameters of interest are the treatment effects
- tests of hypothesis are expressed in terms of the treatment effects

Random model

- factors involved in the experiment are all random
- inference is directed to the population of levels from which the factor levels are randomly selected
- all inferences (parameter estimation and hypothesis testing) are expressed in terms of variance components

Mixed model

- model with both fixed and random factors

Precision vs Accuracy

Precision

- the repeatability of measurements
- used as a measure of the reliability of the experiment
- measured in terms of the variance
- can be increased by increasing the number of samples, skillful grouping of experimental materials and proper selection of treatments

Accuracy

- closeness of the average values of the measurements to the true value
- can be increased through refinement of experimental technique and proper selection of treatments

The 3 basic principles of experimental designs

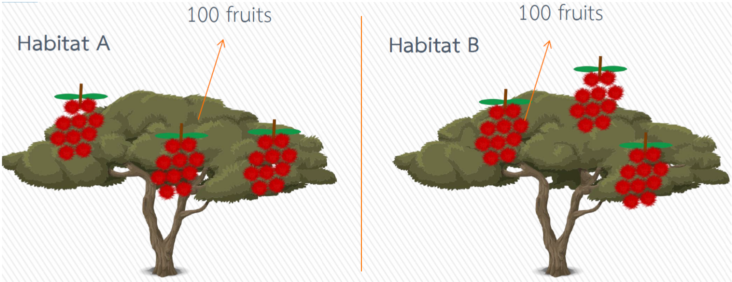
Replication - independent repetition of the basic experiment

- refers to the number of independent experimental units applied with a treatment
- allows us to measure or quantify experimental error
- applying a treatment to a sufficient number of replicate units reduces the effect of uncontrolled variation

The 3 basic principles of experimental designs

The problem of pseudo-replication

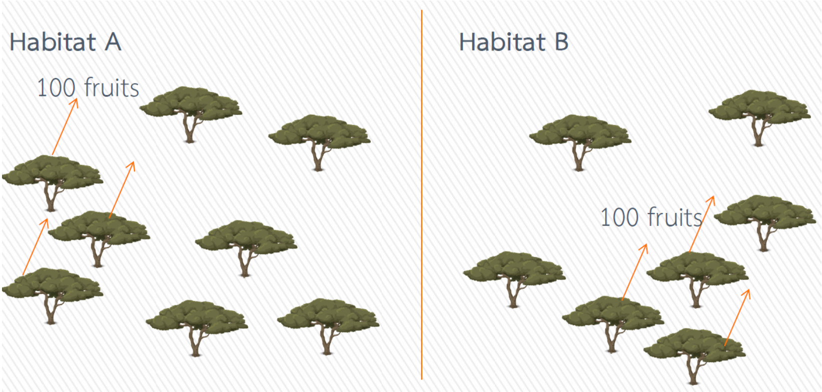
- Pseudo-replication (Hurlbert, 1984) - refers to the use of inferential statistics to test for treatment effects with data from experiments where either treatments are not replicated (though samples may be) or replicates are not statistically independent



The 3 basic principles of experimental designs

The problem of pseudo-replication

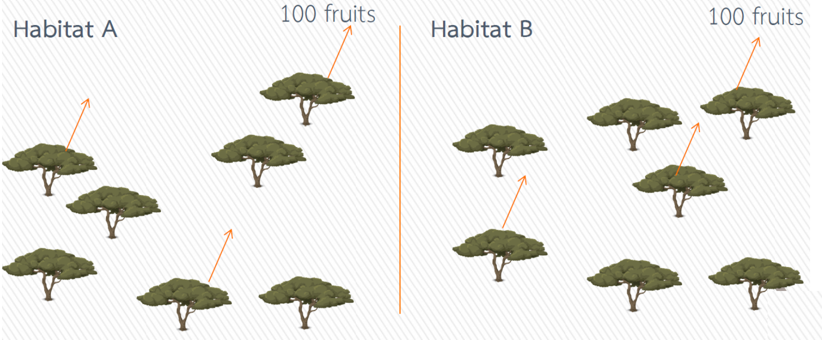
- only tree growing in the same patch were sampled



The 3 basic principles of experimental designs

The problem of pseudo-replication

SOLUTION: Randomly sample trees distributed across each habitat and collect 100 fruits from these trees



The 3 basic principles of experimental designs

How many replicates?

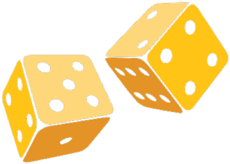
- As a rule of thumb, the number of replicates or blocks should be sufficient to give at least 12 residual degrees of freedom (rdf)
- In more controlled experiments, rdf can be smaller
- It is better to have fewer treatments with sufficient number of replications than having more treatments with few replications
- There are many formulas to approximate the number of replicates per treatment, for example

$$r > \frac{2t^2\sigma^2}{\delta^2}$$

The 3 basic principles of experimental designs

Randomization - random process of assigning treatments to the experimental units such that every treatment has the same chance of being assigned to an experimental unit

- It also means that the experimental units within each treatment represent a random sample from an appropriate population of experimental units
- This ensures that our estimates of population parameters (means, treatment effects, mean squares) are unbiased and our statistical inferences are reliable



The 3 basic principles of experimental designs

Control of extraneous factors

- *Extraneous* or *confounding* factors are factors that may affect the response we wish to measure, but is not included in the study
- These variables may *mask* true treatment effect when not properly taken cared of in the design of the experiment
- If recognizable, their effect can be minimized by proper application of the blocking technique
- *Blocking* - refers to grouping of the experimental units such that units within a block are more or less homogeneous and units between blocks are heterogeneous
- The main purpose of the principle of control is to increase the efficiency of an experimental design by decreasing the experimental error

Structures of a designed experiment

- before an appropriate model and analysis can be constructed for a specific design, all of the factors used in a design of an experiment must be classified as belonging to either the *treatment structure* or the *design structure*

Treatment structure

- consists of the set of treatments, factors, treatment combinations, or populations that the experimenter has selected to study and/or compare
- dictated by the research hypotheses
- blocking factor must not interact with factors in the treatment structure
- one-way, two-way, factorial arrangement treatment structures

Structures of a designed experiment

Design structure

- consists of the grouping of the experimental units into homogeneous groups or blocks
- It is important that experimental units are as uniform as possible prior to applying the treatment structure
- design structure is selected by using all available knowledge of the experimental units and is chosen independently of the treatment structure
- completely randomized design structure, randomized complete block design structure

Types of treatment structures

One-way treatment structure - treatments consist of the levels of a single factor

Two-way treatment structure - the set of treatments was constructed by combining the levels or possibilities of two different factors (no interaction)

Factorial arrangement treatment structure - consists of the set of treatment combinations constructed by combining the levels of two or more factors

- useful when the experimenter is interested on the interaction effect of the two factors

Fractional factorial arrangement treatment structure - consists only a part, or fraction, of the possible treatment combinations in a factorial arrangement treatment structure

Types of design structures

Completely randomized design structure

- all experimental units are assumed to be homogeneous and the experimental units are assigned to the treatments completely at random
- treatments are assigned to an equal number of experimental units, although this is not required

Randomized complete block design structure

- takes care of single source of variation among the experimental units
- number of experimental units within a block is a multiple of the number of treatments
- the complete set of treatments is assigned completely among units within each block

Types of design structures

Latin square design structure

- takes care of two sources of variation among the experimental units
- taken cared of by row blocks and column blocks
- each treatment occurs once and only once in each row block and once and only once in each column block

Incomplete-block design structure

- number of experimental units in a block is less than the number of treatments
- the size of the block cannot accommodate all treatments
- useful especially in experiments with large number of treatments

Other Experimental Designs/Treatment Structures

- Experiments with Nested Factors
- Split-plot Design
- Experiments with repeated measurements

Guidelines for an experimental investigation

- Recognition of and statement of the problem
- Choice of factors, levels and range
- Selection of the response variable
- Choice of experimental design
- Performing the experiment
- Statistical analysis of the data
- Conclusions and recommendations

Roles of the researcher and the statistician

- The investigator and the statistician must work together in designing an experiment (**Design and statistical analysis go hand in hand**)
- The investigator
  - formulates questions of interest
  - decides which comparisons are relevant
  - decide what is a meaningful scientific difference
  - identify the resources available
  - identify peculiarities of the situation
  - decide on the scope of interest
- The statistician:
  - identify the relevant population(s)
  - identify an appropriate probability model
  - translate the question(s) of interest into statistical hypotheses
  - recommend an experimental design

Takeaway

"Block what you can and randomize what you cannot."

— Box, Hunter, and Hunter (1978)