COMMON MISTAKES IN SMALL SAMPLE SIZE BIOLOGICAL EXPERIMENTS AND HOW TO AVOID THEM

Edzard van Santen
IFAS Statistical Consulting Unit & Agronomy Department
University of Florida

THE IFAS STATISTICAL CONSULTING UNIT (IFAS-SCU)

- **Edzard van Santen, Professor and Director**

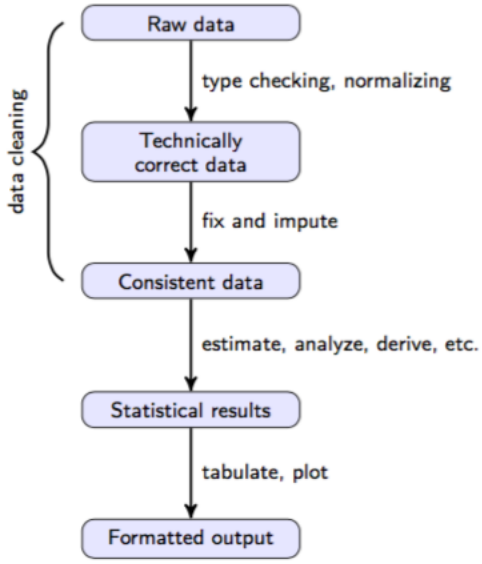
UF | IFAS Research UNIVERSITY of FLORIDA

The IFAS Statistical Consulting Unit

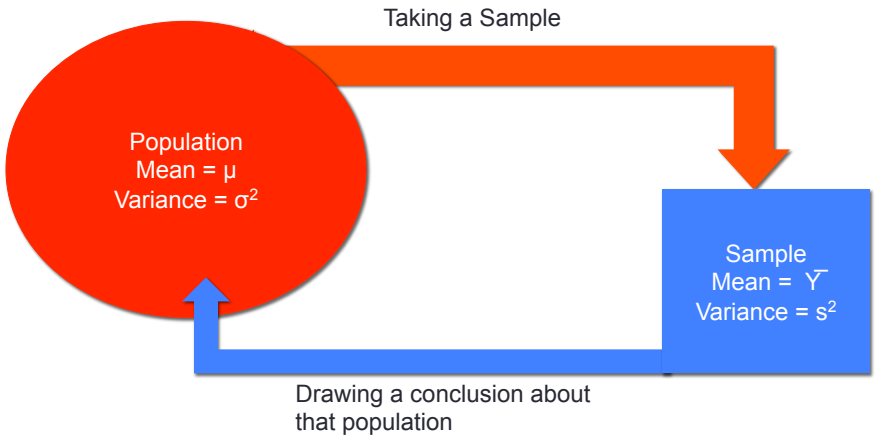
- A **service unit**,

  funded directly by the office of the Senior Vice President for Agriculture and Natural Resources
- Part of the **research infrastructure** of the University of Florida
- Its services are available to anyone within IFAS, both on- and off campus

The Statistical Analysis Value Chain



Key Idea of Statistics



Tales from the Dark Side

- From an application to the AOSCA National Grass Variety Review Board:
  - " 30 plants showing the typical characteristics of the experimental were measured for…."
  - **Background:** this was a cross-pollinated cool season species

QUESTION: Does this strike you as odd?

Tales from the Dark Side

This paper from Hongkong deals with composting hog manure in order to reduce volume and odor and improve fertilizer value. The authors subjected a single batch of manure from a single hog house to three composting treatments (None, Method 1, Method 2). After treatment they mixed each pile (about 300 kg) thoroughly and removed 10 small portions (300 g) from random spots in each pile for laboratory analysis. They compared the treatments using the pooled variation among the 300g-portions within treatments as the error term. The response variable of interest is phosphorous concentration.

QUESTION: Does this strike you as odd?

Tales from the Dark Side

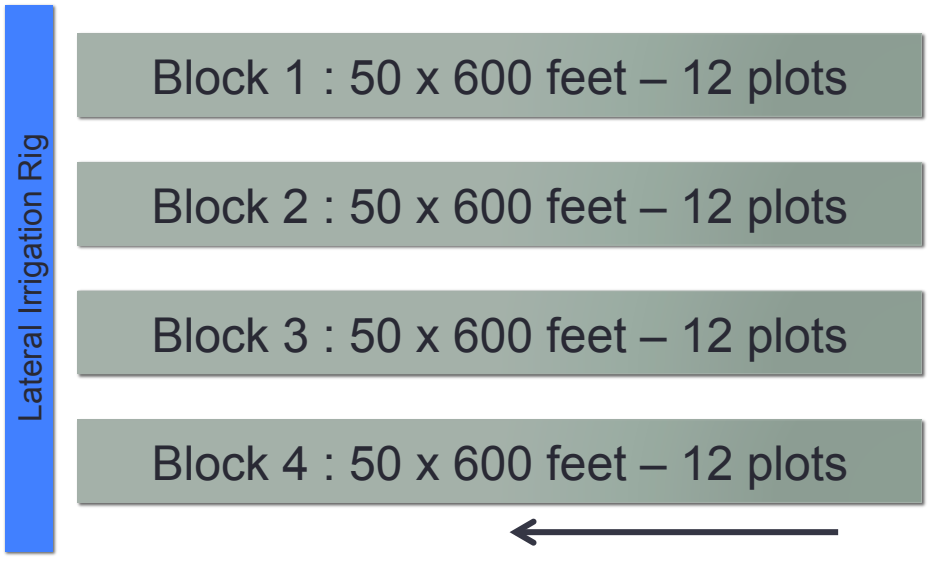
This paper reports results from soil fertility treatments (9 complete blocks * 9 treatments = 81 eu) on Cocoa tree productivity in Benin. The authors sampled the soil by taking 50 small cores throughout each experimental unit. These 50 cores then were thoroughly mixed and a soil sample bag filled with about 250 g of soil for a total of 81 samples for the entire experiment. The researchers later decided to pool the soil samples by treatment in order to reduce the number of samples to be analyzed (and thus cost) from 81 to 9. The samples were sent to a regional soil-testing laboratory, where each of the 9 composite samples was analyzed in duplicate. The authors compared the treatment means using the pooled among lab-replicate variation as the error term.

QUESTION: Does this hurt you as much as it did me?

ARMed and Dangerous

- Based on industry demands, the developers of this software tool for field research enabled a feature that assigns the treatments in the first block (rep) of an experiment in treatment order, rather than randomized.
  - User can override this feature
- What are the consequences?
  - If it is a standard 2-factor herbicide trial the first block will look like a split plot
  - Either inflated Type I error or reduced power but we don't know which.
  - If the experiment had 100 reps, the effect would be minimal

The Long and Skinny



Study types

1. Descriptive (e.g., case-study, naturalistic observation, survey)
2. **Correlational** (e.g., case-control study, observational study)
3. Semi-experimental (e.g., field experiment, quasi-experiment)
4. **Experimental** (experiment with random assignment)
5. Review (literature review, systematic review)
6. Meta-analytic (meta-analysis)

Why do we do experiments?

- Because we are interested in causal relationships
- What does that mean?
  - If you want to see a change in the response you have to make a change in a treatment factor
- The **ceteris paribus** principle
  - all other thing being equal (on par)
- Motto: O R.A. Fisher, in thee have I trusted. Let me never be confounded, …….

Definitions every researcher should know

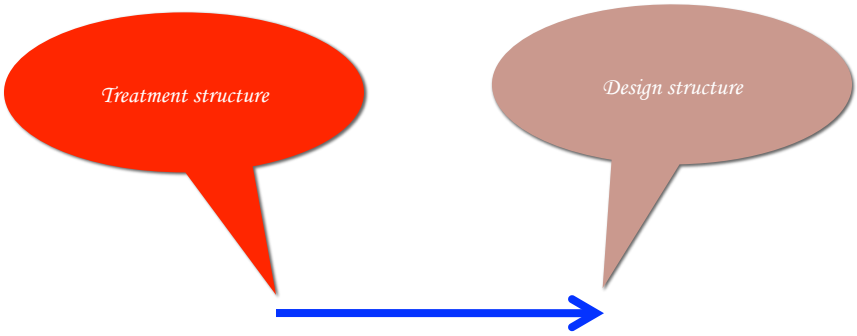
- Experimental Design
- Experiment
  - Replication
  - Randomization
- Experimental Unit
  - Experimental Error
- Sampling Unit
  - Sampling Error
- Treatment Design

Experimental Design is a process

Rules and procedures to

- Plan
- Conduct, and
- analyze an experiment

Experimental Design



Randomization Scheme

Conditions for a valid experiment

- Replication
  - Treatment is evaluated on more than one experimental unit
  - Provides the "noise" estimate against which the "signal" is measured
- Randomization
  - Independent assignment of treatments to experimental units
  - George Casella lists the following benefits
    - **Elimination of systemic bias**
    - **Obtaining a representative sample**
    - **Accounting for extraneous(unknown) confounding variables**

Experimental Unit

- Smallest unit to which a treatment was applied
- **Experimental error**
  - Variance among experimental units treated alike
- **Scale-appropriate error term**
  - The proposed experimental error term should reflect the scale on which an estimate is made.
  - Is a particular "treatment level" truly replicated?
  - Example: Location

Sampling Unit

- Observational unit to within an experimental unit
- **Sampling error**
  - Variance among among observational units within an experimental unit

Treatment Design

- Factor = Type of treatment to be deployed
  - Nrate, Herbicide, Crop Management factor
- Level = Specific state of a factor
  - 50, 100, 150 lbs N/acre

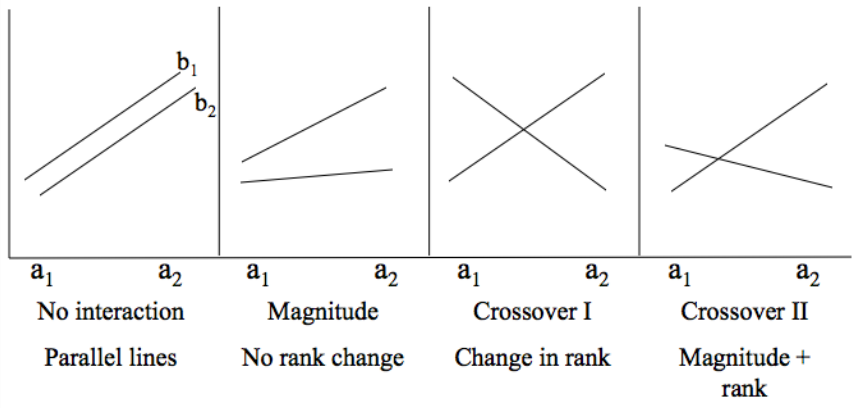
Treatment Design

- Factorial treatment design
  - Two or more factors evaluated jointly, where all levels of one factor occur at all levels of the 2nd factor
  - Simplest design is a 2 x 2 or 2^2 factorial
- Augmented factorial
  - Factorial treatment design plus an "extra" treatment, i.e., a standard control
- Nested treatment design
  - Levels are specific to a given factor
    - Rate in gram a.i. per acre for herbicides

ANALYZE THEM AS YOU RANDOMIZED THEM

➔ Analysis should reflect experiment and treatment design

Interactions



A gene expression experiment

- Field experiment
  - 2 x 2 x 2 factorial (Water, Cultivar, DAP)
  - Experimental Design – RCB (r=3)
    - Split-Block-SP-in-Time restriction on randomization
- Gene expression experiment
  - CRD with technical replication only

- The underlying field experiment design is the proper basis for analysis
- However, The experimental design results in very low df

How can you conduct a better experiment?

- Plan to succeed
  - Can you detect a difference that is biologically meaningful?
- Proper replication
  - Environmental replication will increase scope of inference
- Blocking
  - Think of blocking in both space and time
- Sufficient number treatment levels
  - Estimating a response curve
- Careful experimentation
  - Statistical errors add up, they do not cancel out

Literature

Casler, M.D., Vermerris, W. and Dixon, R.A. 2015 Replication Concepts for Bioenergy Research Experiments Bioenergy. Research 8:1–16

Casler, M.D. 2015. Fundamentals of Experimental Design: Guidelines for Designing Successful Experiments. Agronomy Journal 107: 692-705

Literature

Pearce., S.C. 1988. Analysis of data from agricultural experiments Tropical Agriculture (Trinidad) 65: 2-5

Piepho, H.P., Büchse, A. and Emrich, K. 2003. A Hitchhiker's Guide to Mixed Models for Randomized Experiments. J. Agronomy & Crop Science 189: 310-322

Piepho, H.P., Büchse, A. and Richter, C. 2004. A Mixed Modeling Approach for Randomized Experiments with Repeated Measures. J. Agronomy & Crop Science 190: 230-247

Books

Bowley., S.R. 2015. A Hitchhiker's Guide to Statistics in Biology. Generalized Linear Mixed Models Edition. Plants et al.,Kincardine, Ontario.