Components of an experimental study design

1.1 Study Design: basic concepts
1.2 Factors
1.3 Treatments
1.4 Experimental units
1.5 Sample size and replicates
1.6 Randomization
1.7 Blocking
1.8 Measurements of response variables

Contributors

1.1 Study Design: basic concepts

Usually the goal of a study is to find out the relationships between certain explanatory factors and the response variables. The design of a study thus consists of making decisions on the following:

- The set of explanatory factors.
- The set of response variables.
- The set of treatments.
- The set of experimental units.
- The method of randomization and blocking.
- Sample size and number of replications.
• The outcome measurements on the experimental units - the **response variables**.

### 1.2 Factors

Factors are explanatory variables to be studied in an investigation.

**Examples:**

1. In a study of the effects of colors and prices on sales of cars, the factors being studied are color (qualitative variable) and price (quantitative variable).

2. In an investigation of the effects of education on income, the factor being studied is education level (qualitative but ordinal).

**Factor levels**

Factor levels are the "values" of that factor in an experiment. For example, in the study involving color of cars, the factor car color could have four levels: red, black, blue and grey. In a design involving vaccination, the treatment could have two levels: vaccine and placebo.

**Types of factors**

- **Experimental factors**: levels of the factor are assigned at random to the experimental units.
- **Observational factors**: levels of the factor are characteristic of the experimental units and is not under the control of the investigators.

**Example**: in the "new drug study" (refer to Handout 1), if we are also interested in the effects of age and gender on the recovery rate, then these observational factors; while the treatment (new drug or old drug) is an experimental factor.

### 1.3 Treatments

- In a single factor study, a treatment corresponds to a factor level; thus the number of treatments equals the number of different factor levels of that factor.

- In a multi-factor study, a treatment corresponds to a *combination of factor levels across different factors*; thus the number of all possible treatments is the product of the number of factor levels of different factors.

**Examples:**

- In the study of effects of education on income, each education level is a treatment (high school, college, advanced degree, etc).

- In the study of effects of race and gender on income, each combination of race and gender is a treatment (Asian female; Hispanic male, etc).

**Exercise**: How many different treatments are there for the above examples?
Choice of treatments

Choice of treatments depends on the choice of: (i) the factors (which are the important factors);
(ii) levels of each factor.

- For **qualitative factors** the levels are usually indicated by the nature of the factor.

Example: gender has two levels: female and male

- For **quantitative factors** the choice of levels reflects the type of trend expected by the investigator.

Example: linear trend implies two levels; quadratic trend implies three levels. Usually 3 to 4 equally spaced levels are sufficient.

- The range of the levels is also crucial. Usually prior knowledge is required for an effective choice of factors and treatments (refer to the "quick bread volume" example on page 650).

1.4 Experimental units

- An experimental unit is the smallest unit of experimental material to which a treatment can be assigned.

*Example:* In a study of two retirement systems involving the 10 UC schools, we could ask if the basic unit should be an individual employee, a department, or a University.

*Answer:* The basic unit should be an entire University for practical feasibility.

- **Representativeness:** the experimental units should be representative of the population about which a conclusion is going to be drawn.

*Example:* A study conducted surveys among 5,000 US college students, and found out that about 20% of them had uses marijuana at least once. If the goal of the study is the drug usage among Americans aging from 18 to 22, is this a good design?

- Choosing a representative set of experimental units which fits the purpose of your study is important.

1.5 Sample size and replicates

Loosely speaking, sample size is the number of experimental units in the study.

- Sample size is usually determined by the trade-off between statistical considerations such as power of tests, precision of estimations, and the availability of resources such as money, time, man power, technology etc.

- In general, the larger the sample size, the better it is for statistical inference; however, the costlier is the study.

- An important consideration in an experimental design is *how to assess power or precision as a function of the sample size* (sample size planning/power calculation)?
Replicates

For many designed studies, the sample size is an integer multiple of the total number of treatments. This integer is the number of times each treatment being repeated and one complete repetition of all treatments (under similar experimental conditions) is called a complete replicate of the experiment.

- **Example:** In a study of baking temperature on the volume of quick bread prepared from a package mix, four oven temperatures: low, medium, high and very high were tested by randomly assigning each temperature to 5 package mixes (all of the same brand). Thus the sample size is $20 (= 4 \times 5)$, the number of treatments is 4 (4 levels of temperatures) and there are 5 complete replicates of the experiment.

Why replicates?

When a treatment is repeated under the same experimental conditions, any difference in the response from prior responses for the same treatment is due to random errors. Thus replication provides us some information about random errors. If the variation in random errors is relatively small compared to the total variation in the response, we would have evidence for treatment effect.

1.6 Randomization

- Randomization tends to average out between treatments whatever systematic effects may be present, apparent or hidden, so that the comparison between treatments measure only the pure treatment effect.
- Randomization is necessary not only for the assignment of treatments to experimental units, but also for other stages of the experiment, where systematic errors may be present.

**Example:** In a study of light effects on plant growth rate, two treatments are considered: brighter environment vs. darker environment. 100 plants are randomly assigned to each treatment (all genetically identical). However, there is only one growth chamber which can grow 20 plants at one time. Therefore the 200 plants need to be grown in 10 different time slots.

In addition to randomizing the treatments, it is important to randomize the time slots also. This is because, the conditions of the growth chamber (such as humidity, temperature) might change over time. Therefore, growing all plants with brighter light treatment in the first 5 time slots and then growing all plants with darker light treatment in the last 5 time slots is not a good design.

1.7 Blocking

In a **blocked experiment**, heterogenous experimental units (with known sources of heterogeneity) are divided into homogenous subgroups, called blocks, and separate randomized experiments are conducted within each block.

- **Example:** in a study of Vitamin C on cold prevention, 1000 children were recruited. Half of them were randomly chosen and were given Vitamin C in their diet and the other half got placebos. At the end of the study, the number of colds contracted by each child was recorded. (This is an example of a complete randomized design (CRD).)
- If we know (or have sufficient reason to believe) that gender may also influence the incidence of cold, then a more efficient way to conduct the study is through blocking on gender: 500 girls and 500 boys were recruited. Among the girls, 250 were randomly chosen and given Vitamin C and the other 250 were given placebo. Same is done for the
500 boys. (This is an example of a randomized block design (RCBD).)

- By blocking, one removes the source of variation due to potential confounding factors (here it is gender), and thus improves the efficiency of the inference of treatment effect (here it is Vitamin C).
- Randomization alone (as in CRD) does not assure that the same number of girls and boys will receive each treatment. Thus if there is a difference of cold incidence rate between genders, observed differences between treatment groups maybe observed even if there is indeed no treatment effect.

1.8 Measurements of response variables

The issue of measurement bias arises due to unrecognizable differences in the evaluation process.

Example: The knowledge of the treatment of a patient may influence the judgement of the doctor. The source of measurement bias can be reduced to concealing the treatment assignment to both the subject and the evaluator (double-blind).

Contributors

- Anirudh Kandada (UCD)