Estimators and estimates I

- For a population parameter θ to be able to place it within an interval, $a \leq \theta \leq b$, at a certain confidence level $(1-\alpha)$ we need the following three items:
  - The point estimate $\theta^\star$.
  - A sample statistic that relates $\theta^\star$ to the population parameter θ.
  - The sampling distribution of the statistic.
- We have $P(\bar{X} - z_{\frac{\alpha}{2}}\sigma_{\bar{x}} \leq \mu \leq \bar{X} + z_{\frac{\alpha}{2}}\sigma_{\bar{x}}) = 1 - \alpha$. Note that this is an estimator.
- If we substitute specific values for $\bar{X}$, say $\bar{x}$, corresponding to a specific sample the probability statement does not hold for the specific values calculated from the sample.
- However, what we can say is that if we draw samples of size N a large number of times then $1-\alpha$ fraction of the times the interval $\bar{x} - z_{\frac{\alpha}{2}}\sigma_{\bar{x}} \leq \mu \leq \bar{x} + z_{\frac{\alpha}{2}}\sigma_{\bar{x}}$ will contain μ.

Experiment design

- A **factor** is an independent variable that is controlled.
- The **levels** of a factor is the number of distinct values that a factor can take.
- In many experiments the number of levels a factor can take is finite and small - very often two.

Factorial design I

- Let ℓ be the number of levels of each factor and f the number of factors. Then ℓ^f will be the number of distinct experimental conditions.
- More generally, if $l_1, \ldots, l_f$ are the levels for each factor i (l_i is the number of levels for the ith factor) then the number of distinct experimental conditions is $l_1 \times l_2 \times \ldots \times l_f$.
- In what follows the number of levels in each factor is assumed to be equal and denoted by ℓ.
- The number of actual runs or equivalently the number of distinct experimental conditions that are run will be denoted by r.
- A full factorial design has to run each experimental condition. So for such a design $r = \ell^f$ or $r = l_1 \times \ldots \times l_f$.

Factorial design II

- The process of running each experimental condition multiple times (i.e. > 1) is called **replication** and each individual run of a condition in a replication is called a **replicate**. Replication is desirable since it can average out chance or nuisance factors but is not always possible.
- One key factor in running experiments correctly is **randomization**. Randomization involves i) randomly ordering the experimental conditions and ii) randomly assigning subjects to experimental conditions (where necessary).
- Randomization averages out nuisance variables and uncontrollable variables. It also removes biases when subjects are assigned to conditions.

Fractional factorial designs I

- When ℓ^f is large a fully factorial design is not possible. If r, where $r < \ell^f$, is the number of experimental conditions that can be run then we have a **fractional factorial design**.
- In a fractional factorial design it is necessary to choose the experimental conditions that will be run properly so that the maximum information can be obtained from the experiment. The example below illustrates this. Consider $\ell = 2$, $f = 3$, $r = 4$. Let the factors be f_1, f_2, f_3 and let the two levels of each factor be 0 and 1. Then we have the following 8 distinct experimental conditions.

No.	1	2	3	4	5	6	7	8
f_1	0	0	0	0	1	1	1	1
f_2	0	0	1	1	0	0	1	1
f_3	0	1	0	1	0	1	0	1

Fractional factorial designs II

- There are 70 possible ways to choose 4 runs. Intuitively, we see that many of them give less information - for example runs 1 to 4 or 5 to 8 have only one level for f_1. To get maximum information factors should occur equally at the two levels in the 4 runs - that is twice each at levels 0 and 1. There are many such possibilities. Which is chosen depends on what is known about the domain of the experiment and any known apriori interaction between the factors.
- For example here are a few possible choices where each level for each factor occurs exactly twice (the numbers are the column numbers from the experimental conditions table earlier):

  1,3,6,8; 2,3,5,8; 2,4,5,7.

One or more independent variables are categorial

- Many experiments create groups of subjects who are treated differently. The goal is to understand how the different treatments affect the response variable(s) of interest.
- The simplest and most common case is where there are two groups. The first is the treatment group and the second is the control group. This design is common in drug or other therapeutic trials. The aim is to check if the treatment has a statistically significant effect.
- A second example. Suppose there are 4 different proposals on how to improve math teaching for primary school children where improvement is measured by a score on a common math test administered to all students before and after the intervention. This experiment has 4 groups and we would like to know which interventions have a statistically significant effect.
- In all the above one (or more) independent variables can be categorial/ nominal based on which the groups are created.