Terminology

- **Response variable** (dependent variable): the variable of interest in the study
- **Factors** : the variables that may have an effect of the response variable
  - quantitative if measured on a numerical scale
  - qualitative otherwise
- **Factor Levels**: the values of the factors utilized in the experiment
- **Treatments**: the factor-level combinations utilized.
- **Experimental Units** (subjects): the objects on which the factors are measured or observed.

Therefore:

- A *designed experiment* is one for which the analyst or experimenter **controls** the specification of treatments and the method of assigning units to treatments.
- An *observational experiment* or study is one for which the analyst simply **observes** the treatments and response on a sample of experimental units.

Example: Birthweight study.

- **Response:** Birthweight (g)
- **Factor:** Pre-natal treatment group
- **Factor levels:** GROUP A or GROUP B

that is, we have a single factor with two factor levels.

Example: SAT scores.

The SAT scores of female and male students in four schools are to be compared.

- **Response:** SAT score
- **Factors:** SEX and SCHOOL (both qualitative)
- **Factor levels:**
  - SEX: Female and Male
  - SCHOOL: A,B,C,D

that is, we have a two factors, SEX with two factor levels and SCHOOL with four factor levels. There are 8 possible treatments:

$$(F, A), (F, B), (F, C), (F, D), (M, A), (M, B), (M, C), (M, D)$$

Example: Pain Relief.

Different pain relief remedies are to be compared : factors are

- REMEDY (quantitative/qualitative, 3 levels)
  - Dose level 0
  - Dose level 1
  - Dose level 2
- AGE GROUP (quantitative/qualitative, 4 levels)
  - 0-16 years
  - 17-40 years
  - 41-65 years
  - 66 years and over
- SEX (qualitative, 2 levels)
  - Female
  - Male

A total of $3 \times 4 \times 2 = 24$ possible treatment combinations; REMEDY is the only factor that can be assigned by the analyst.

Completely Randomized Design

A *completely randomized design* (CRD) is a design for which treatments are randomly assigned to experimental units, or in which random samples of experimental units are selected for each treatment.

The term can be applied to both experimental and observational studies. For example,

- if the treatments are FEMALE/MALE for the factor SEX, a CRD draws independent samples of FEMALES and MALES for the two treatment groups.
- if the treatments are DOSE 0/DOSE 1, a CRD assigns experimental units independently to the two treatment groups at random.

Statistical Objectives

The experimental units assigned to different treatments (factor-level combinations) form

independent samples

from

different populations

in a CRD.

We wish to **compare** treatments: specifically, we wish to compare the treatment means.

A Multiple Group Comparison of Means !

Suppose that there are k treatments:

TMT 1	Mean μ_1
TMT 2	Mean μ_2
$\vdots$	$\vdots$
TMT k	Mean μ_k

We wish to test the hypotheses

H_0 : $\mu_1 = \mu_2 = \cdots = \mu_k$

H_a : At least two of the k treatment means are different

How do we do this ?

What is the relevant test statistic ?