STAT 332: Sampling and Experimental Design

Professor Riley Metzger
LaTeXer Iris Jiang

Spring 2020

Contents

1 PPDAC

Problem, Plan, Data, Analysis, Conclusion

1.1 Problem

Define the proble:

- Target Population (T.P.): The group of units referred to in the problem step
- Response: The answer provided by the T.P. to the problem
- Attribute: statistic of the response

Example 1.1. What is the average grade of students in STAT 101?

Solution.

- T.P.: All STAT 101 students
- Response: Grade of a STAT 101 student
- Attribute: Average grade

□

1.2 Plan

How?

- Study population (S.P.): The set of unites you can study

  Example 1.2. Problem: Does a drug reduce hair loss

  Solution. You can not use untested drug directly on people out of ethical concerns

  T.P.: People

  S.P: Mice □

- Sample: A subset of the study population

1.3 Data

Collect the data, according to the plan.

1.4 Analysis

Analyse the data.

1.5 Conclusion

Refers back to the problem.

1.6 Errors

- Study Error: The attribute of the T.P. differs from the parameter of the S.P.

  Example 1.3. $a(T.P.) - \mu$

- Sample Error: The parameter differs from the sample statistic (estimate).

  Example 1.4. $\mu - \bar{x}$

- Measurement Error: The difference between what we want to calculate and what we do calculate.

2 Models

Definition 2.1 (Model)**.** A model relates a parameter to a response.

2.1 Model I

$$Y_j = \mu + R_j,\ R_j \sim N(0, \sigma^2)$$

- y_j: The response of unit j, it is random.
- μ: S.P. mean, it is not random and it is unknown
- R_j: The distribution of responses about μ

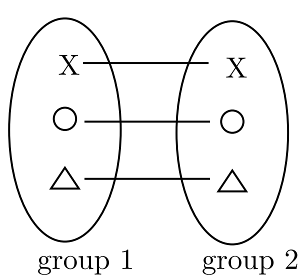
Note.

1. *R_j's are always independent.*
2. *Gaus's Theorem: Any Linear combination of normal R.V.s is normal*
3. $Y_j \sim N(\mu, \sigma^2)$,
$$E(Y_j) = E(\mu + R_j) = E(\mu) + \mu + 0 = \mu$$
$$V(Y_j) = V(\mu + R_j) = V(R_j) = \sigma^2$$

Example 2.1. Average grade of STAT 101: $Y_j = \mu + R_j,\ R_j \sim N(0, \sigma^2)$

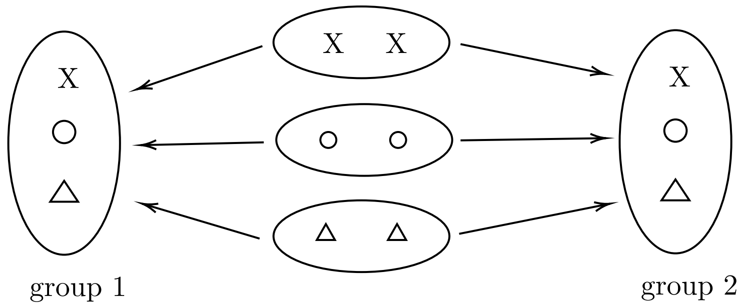
2.2 Independent vs. Dependent Groups

Definition 2.2 (Dependent)**.** We randomly select one group and we find a match, having the same explanatory variates, for each unit of the first group.

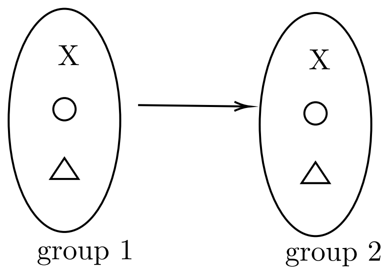


2.2.1 Ways of Creating Dependency

- Twins



- Reuse



Definition 2.3 (Independent)**.** Are formed when we select units at random from mutually exclusive groups.

- No relationship between chosen groups

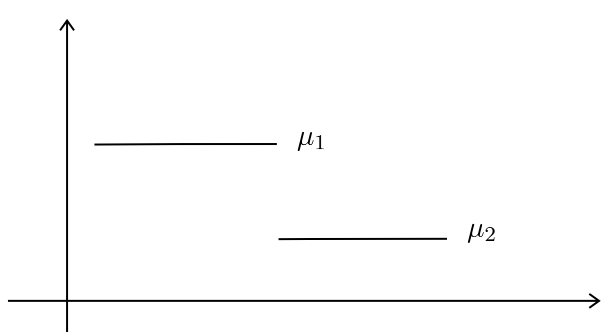
  Example 2.2. Broken parts and non-broken parts

2.3 Model 2A

Independent groups where we assume the groups have the same standard deviation.

$$Y_{ij} = \mu_i + R_{ij},\ R_{ij} \sim (0, \sigma^2)$$

- Y_{ij}: Response of unit j in group i
- μ_i: Mean for group i; not random; unknown
- R_{ij}: The distribution of responses about μ_i



2.4 Model 2B

Independent groups but $\sigma_1 \neq \sigma_2$

$$Y_{ij} = \mu_i + R_{ij}, R_{ij} \sim N(0, \sigma_i^2)$$