

§ 5.1 FUNDAMENTALS OF STATISTICS

Def: STATISTICS IS THE SCIENCE OF COLLECTING, ORGANIZING, & INTERPRETING DATA. THE DATA ITSELF ARE ALSO CALLED STATISTICS.

Def: THE POPULATION IS THE COMPLETE SET OF THINGS BEING STUDIED.
THE SAMPLE IS THE SUBSET OF THE POPULATION FROM WHICH RAW DATA (INDIVIDUAL MEASUREMENTS) ARE OBTAINED.

POPULATION PARAMETERS ARE SPECIFIC CHARACTERISTICS OF THE POPULATION THAT WE WISH TO KNOW.

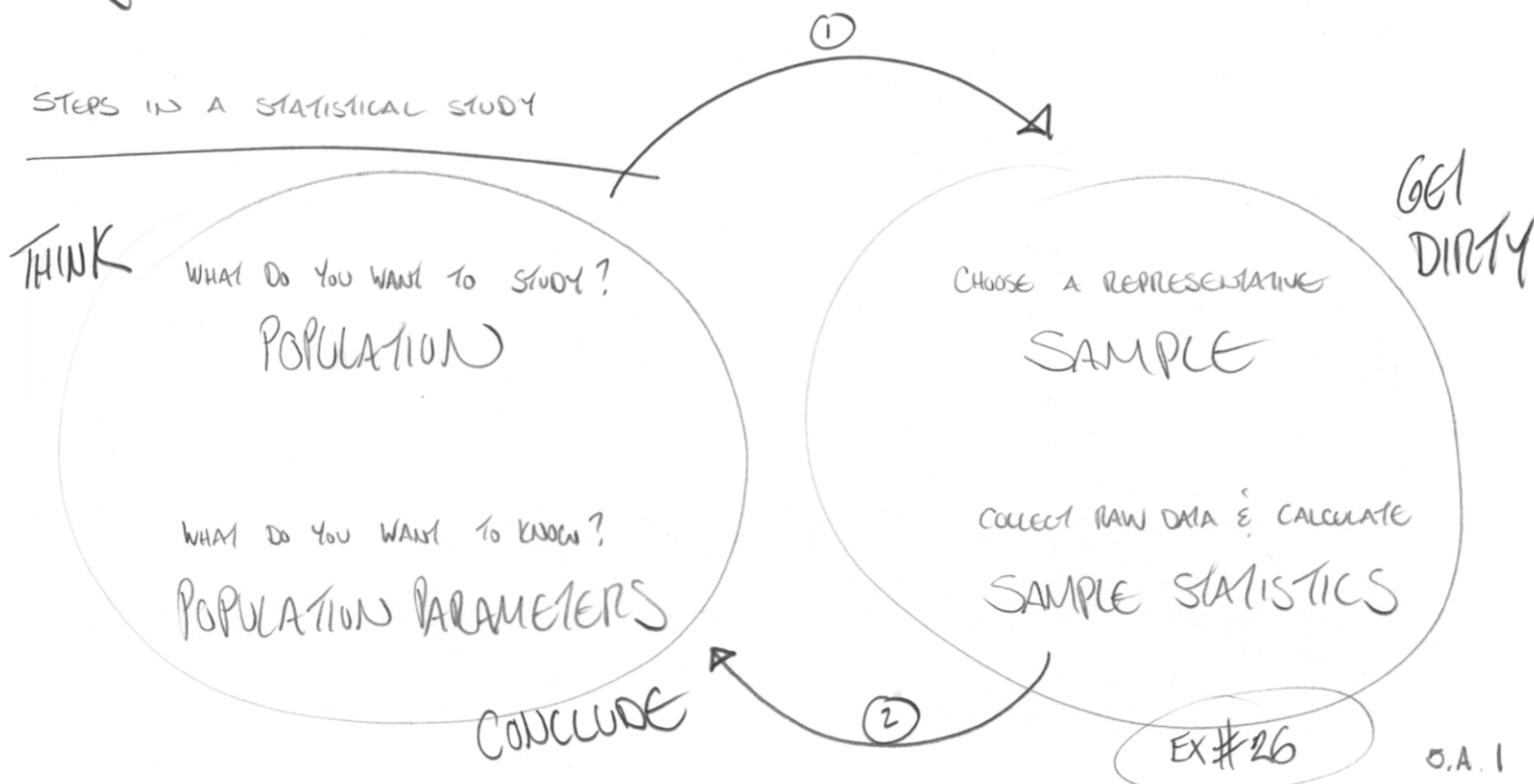
SAMPLE STATISTICS ARE SPECIFIC CHARACTERISTICS OF THE SAMPLE THAT WE DO KNOW (AND HOPE APPROXIMATE THE POPULATION PARAMETERS WELL).

e.g. NIELSON RATINGS.

e.g. INTERNET POLLS / SURVEYS

e.g. GALLUP POLLS

STEPS IN A STATISTICAL STUDY



Def THE CHARACTERISTICS OF THE MEMBERS OF A REPRESENTATIVE SAMPLE
MATCH THOSE OF THE POPULATION.

e.g. AVERAGE FAMILY SIZE → COSTCO, DAY CARE
→ NIGHT CLUB

EX #28

e.g. AVERAGE COMMUTE TIME → PLATFORM OF 1 TRAIN
→ PLATFORM OF LIRR

SAMPLING METHODS

SIMPLE RANDOM SAMPLE : EVERY SAMPLE OF SIZE n HAS EQUAL CHANCE
OF BEING SELECTED.

SYSTEMATIC SAMPLE : SELECT EVERY n^{TH} MEMBER OF THE POPULATION

CONVENIENCE SAMPLE : SELECT A GROUP THAT IS EASY TO SINGLE OUT
FOR STUDY TO BE THE SAMPLE.

STRATIFIED SAMPLE : BREAK POPULATION INTO SUBGROUPS (STRATA)
THEN SELECT S.R.S.'S FROM EACH.

USED WHEN CONCERNED ABOUT DIFFERENCES BETWEEN

SUBGROUPS - e.g. MEN VS. WOMEN

e.g. 7TH VS. 8TH GRADERS

ALL THAT MATTERS IS WHETHER A SAMPLE IS REPRESENTATIVE.

WE CAN'T REALLY KNOW, BUT LARGER SAMPLES ARE MORE LIKELY
TO BE REPRESENTATIVE.

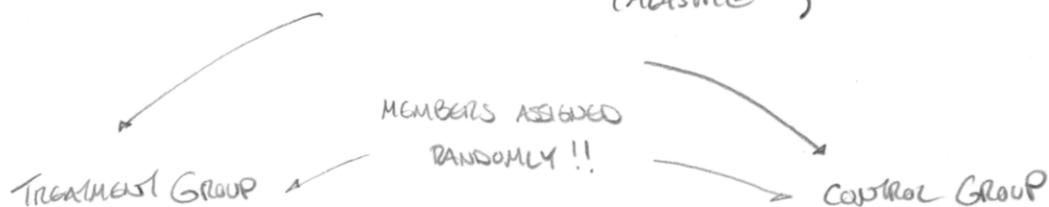
EX 31, 34

Def: A STATISTICAL STUDY SUFFERS FROM BIAS IF ITS DESIGN OR CONDUCT TENDS TO FAVOR CERTAIN RESULTS.

TYPES OF STATISTICAL STUDY

Def: OBSERVATIONAL STUDY - RESEARCHERS OBSERVE/MEASURE CHARACTERISTICS OF THE SAMPLE W/O INTERFERING

EXPERIMENT - RESEARCHERS APPLY TREATMENT TO SOME OR ALL OF SAMPLE & LOOK AT THE EFFECTS.
(MEASURE)



EVERYTHING THE SAME EXCEPT THE VARIABLE BEING STUDIED !!

PLACEBO EFFECT

BLINDING

(NOT INFORMING PEOPLE WHETHER THEY ARE IN TREATMENT OR CONTROL)

SINGLE-BLIND : PARTICIPANTS DON'T KNOW
RESEARCHERS KNOW.

DOUBLE-BLIND : PARTICIPANTS DON'T KNOW
RESEARCHERS DON'T KNOW (TILL THE END)

CASE CONTROL STUDIES

↳ AN OBSERVATIONAL STUDY THAT RESEMBLES AN EXPERIMENT BECAUSE THE SAMPLE DIVIDES (NATURALLY) INTO TWO (OR MORE) GROUPS. THOSE WHO ENGAGE IN BEHAVIOUR BEING STUDIED FORM CASES (LIKE TREATMENT GROUP). THOSE WHO DON'T FORM CONTROL.

e.g. DO SEAT BELTS SAVE LIVES?

e.g. ALCOHOL/DRUGS DURING PREGNANCY.

Surveys & opinion Polls

MARGIN OF ERROR IS USED TO DESCRIBE CONFIDENCE INTERVAL
(95%)

e.g. 69% OF AMERICANS WANT STRICTER GUN CONTROL,
WITH MARGIN OF ERROR OF 2%.

⇒ THE TRUE PERCENTAGE OF AMERICANS ... IS BETWEEN

67% & 71%, AND IF THE POLL WERE CONDUCTED

20 TIMES, 19 OUT OF 20 WOULD PRODUCE A CONFIDENCE

INTERVAL CONTAINING THE TRUE POP. PARAM.

} 95%
CONFIDENCE