

Ch. 9 – 11 – Gathering Data

Def'n: An observational study is a study where a researcher observes characteristics of subjects in samples from populations of interest.

A retrospective study is an observational study in which subjects are selected and then their previous conditions/behaviours are determined.

A prospective study is an observational study in which subjects are followed to observe future outcomes. No treatments are deliberately applied.

An experiment is a study where a researcher applies different treatments to different subjects and observes the outcomes. A controlled clinical trial is a type of experiment.

Drawing conclusions:

1. Infer to a larger population (Population)
2. Factor causes change in response (Causal)

		Random sampling?	
		Yes	No
Random Assignment?	Yes	Both inferences are possible	Causal inferences
	No	Population inferences	Neither is possible

Both types of study allow for population inferences, but only a properly designed (and *randomized*) experiment allows for causal inferences to be valid. Experiments are not always feasible.

Sampling

Def'n: A sampling frame is the list of subjects in the population from which the sample is taken.

A random sample is a sample drawn in such a way that each element of the population has a chance of being selected. If chances are all the same → SRS of size n

Ex9.1) A deck of cards: picking a card is a simple random sample. Moreover, placing the card back in the deck is a sample *with replacement*. Otherwise, there is sampling *without replacement*.

Other sampling methods:

- *stratified random sample*: divide population into strata, SRS from each stratum.
- *cluster random sample*: divide population into large # of clusters, select SRS of clusters.
- *systematic sample*: select individuals systematically from a sampling frame.
- *convenience sample*: select individuals who are conveniently available.
- *voluntary response sample*: collect data from individuals who volunteer their responses.

Since the first two use simple random sampling at different points of the data collection process, it is best to compare SRS to stratified and cluster random sampling.

SRS: Advantage: sample tends to be a good reflection of population

Disadv.: sample may not reflect well if sample size is not large enough

Stratified RS: Adv.: ensures enough subjects in each group to compare

Disadv.: must have a sampling frame, must know how data separates into strata,
more costly since each stratum must be used

Cluster RS: Adv.: do not need a sampling frame, less expensive to implement

Disadv.: still same problems as SRS if size not large enough

Bias:

- undercoverage: samples differ due to systematic exclusion of part of the population

Ex9.2) excluding French citizens with a phone survey in English

- response bias: samples differ because of method of observation

Ex9.3) if 4 of 5 dentists recommend Trident, is gum bad for you?

- nonresponse bias: samples differ due to unobtainable data

Ex9.4) sending surveys to Terrans &
extra-terrestrials (who don't send the surveys back in non-Glorpian style)

More terms to know

Def'n: A control group is an experimental group that receives no treatment.

A placebo is identical to a treatment but definitively has no effect.

A single-blind experiment is where the subjects are unaware of which treatment is received, but investigator knows. Conversely, also possible for investigator to be unaware while subject knows.

A double-blind experiment has both subjects and investigators unaware of which treatment is received.

A design of an experiment is the overall plan for conducting the experiment.

A factor is a categorical explanatory variable. "Extraneous" factors may exist.

Four principles of experimental design:

Direct Control: hold or fix these factors at a constant level

Randomization: random assignment of treatments to remove effect of particular condition

Replication: repeating treatment enough for adequate sample size and more confidence in results

Blocking: arrange groups by these factors, apply all treatments inside each block