Statistics Misuse, Experimental Design

PROF. DR. AKIN PALA

Types of Statistics, Cross-Section Versus Time-Series

- Based on the time over which they are collected, data can be classified as either
- Cross-section data
  - information on different members/elements of a population (subjects) or sample for the *same period of time*
    - *Example: Reaction time of 200 people after drinking a 250 ml bottle of Cola*
- Repeated measures (time-series data)
  - Data collected on the same subject for the same variable at different points in time
    - Example: *Reaction time of 200 people after drinking 1 bottle of Cola, the time after the 2nd and the time after the 3rd bottles in an hour.*

Repeated Measures (Time Series)

- The first Cola may have an effect on the 2nd and the 3rd, the 2nd bottle may have an effect on the 3rd
  - This is called a carry over effect and has to be taken into account in statistical calculations using covariance structures.
- If a newborn baby mouse is on a low level of nutrition, it may have lower weight or IQ later in life
  - How do you measure IQ of mice?
  - In a maze ☺

Always look at the methodology

- Saanen goats have an average of 1.7 kids
- This means “In a representative sample of 10 Saanen goats, we can expect 17 kids”
- How about if an average Saanen family has .90 kids?
- If they have a tendency for twinning how did that happen?

Misuse

- Misuse of statistics can produce subtle, but serious errors in description and interpretation
- Subtle because experienced professionals can make such errors.
  - Heavily depending on software may lead to inaccurate results. The output must be checked before publishing.
- Serious because they can lead to devastating decision errors.
  - Social policy, medical practice, central university exam results, drug usage, battery drain, which TV shows get more commercials and the reliability of biology experiments all rely on the proper use of statistics.

Misuse

- Even when statistics is correctly used, the results can be difficult to interpret.
  - 3x3 interactions, main effects, multivariate results etc.
- The statistical significance, measuring ratio of differences to random variation —may or may not agree with an intuitive sense of its significance.
- You will be given the statistical literacy so you can understand
  - Everyday statistics
  - Experimental results

Misuse, General Perception

- The general perception is that statistical knowledge is generally misused by interpreting only the favorable data.
- A mistrust and misunderstanding of statistics is associated with the quotation, "There are three kinds of lies: lies, damned lies, and statistics".
  - “They are lying using statistics”
- Misuse of statistics can be involuntary, unconscious or intentional.

Misuse

- Ways to avoid misuse of statistics include using proper diagrams and avoiding bias.
- Misuse can occur when conclusions are over generalized and claimed to be representative of more than they really are, often by either deliberately or unconsciously overlooking sampling bias.
- Bar graphs are arguably the easiest diagrams to use and understand, and they can be made either by hand or with simple computer programs.

Misuse

- Unfortunately, most people do not look for bias or errors, so they are not noticed.
- Thus, people may often believe that something is true even if it is not well represented.
- To make data gathered from statistics believable and accurate, the sample taken must be representative of the whole.

Misuse

A series of questions to assist in statistics :

- Who says so? (Does he/she have an axe to grind?)
- How does he/she know? (Does he/she have the resources to know the facts?)
- What's missing? (Does he/she give us a complete picture?)
- Did someone change the subject? (Does he/she offer us the right answer to the wrong problem?)
- Does it make sense? (Is his/her conclusion logical and consistent with what we already know?)
  - It is very hard to convince people when you find an outlier.

Misuse, Misinterpretation: correlation

- The concept of correlation can cause confusion.
- Statistical analysis of a data set often reveals that two variables (properties) of the population under consideration tend to vary together, as if they were connected.
  - For example, a study of annual income that also looks at the age of death might find that poor people tend to have shorter lives than people doing well.
  - The two variables are said to be correlated; however, they may or may not be the cause of one another.

Misuse, Misinterpretation: correlation

- The correlation phenomena could be caused by a third, previously unconsidered phenomenon, called a confounding variable.
- There is no way to immediately infer the existence of a causal relationship between the two variables.
  - Correlation does not imply causation.
- A correlation between age and height in children is clear
- A correlation between mood and health in people is not
  - Does improved mood lead to improved health, or does good health lead to good mood, or both? Or does some other factor underlie both? A correlation can be taken as evidence for a possible causal relationship, but cannot indicate what the causal relationship, if any, might be.

Misuse, Misinterpretation: correlation

- An increase in Imams' salaries and an increase in wine sales may be correlated. This may not necessarily mean the imam's started drinking as their salary increased.
- There could be due to a third, confounding factor:
  - Taxation in economy,
  - A policy in encouraging imams to work harder in the face of a fast increasing # of tourists, who happened to drink a lot wine.

Misuse, Misinterpretation: correlation

- As ice cream sales increase, the rate of drowning deaths increases sharply.
  - Therefore, ice cream consumption causes drowning.
- Ice cream is sold during the hot summer months and people tend to go swimming in summer.
- The drowning deaths are caused by more exposure to water-based activities, not ice cream.
- Joke: Temel ripped wings of a fly and yelled: fly
  - It did not fly, so he concluded ripping the wings causes deafness in flies.

Misuse, Misinterpretation: correlation

- Young children who sleep with the light on develop myopia (distant objects are out of focus) later in life.
- Therefore, sleeping with the light on causes myopia.
  - University of Pennsylvania Medical Center. Published in Nature by Quinn, Graham E.; Shin, Chai H.; Maguire, Maureen G.; Stone, Richard A. (May 1999). "Myopia and ambient lighting at night". Nature 399 (6732): 113–4.
- The study received much coverage in the popular press.
- A later study did not find that infants sleeping with the light on caused the development of myopia.
- It found a strong link between parental myopia and the development of child myopia, also noting that myopic parents were more likely to leave a light on in their children's bedroom.
- In this case, the cause of both conditions is parental myopia, and the above-stated conclusion is false.

Confoundation

- The confounding variable problem: X and Y may be correlated, not because there is causal relationship between them, but because both depend on a third variable Z.
- Z is called a confounding factor.
- X and Y may be two enzymes, and Z may be the gene.
- The enzymes are not affecting each other, they are correlated because they are expressed by the same gene.
  - Gene regulation, alternative splicing:
  - A single gene can produce multiple sequences of amino acids, depending on which exons are included.

Confoundation

- There are two reasons:
- **Pleiotropy** occurs when one gene influences two or more seemingly unrelated phenotypic traits
- Another example: phenylketonuria is caused by one gene defect, but it affects multiple systems.
  - A mutation in a **pleiotropic** gene may affect some or all traits simultaneously.
- **Genetic linkage** occurs when there is a tendency of certain genes to be inherited together.

Confoundation

- Comparing brunette people from Izmir with blondes from Canakkale
- If blondes from Canakkale performed better in the statistics final exam
  - Is it because they are blonde?
  - Is it because they are from Canakkale?
- Consulting before designing experiments is important
- You need
  - blondes from Canakkale
  - blondes from Izmir
  - brunettes from Canakkale
  - brunettes from Izmir

Success in Statistics

- Success in this introductory statistics course typically requires more common sense than mathematical expertise.
- This section is designed to illustrate how common sense is used when we think critically about data and statistics.

Misuse, Bad samples

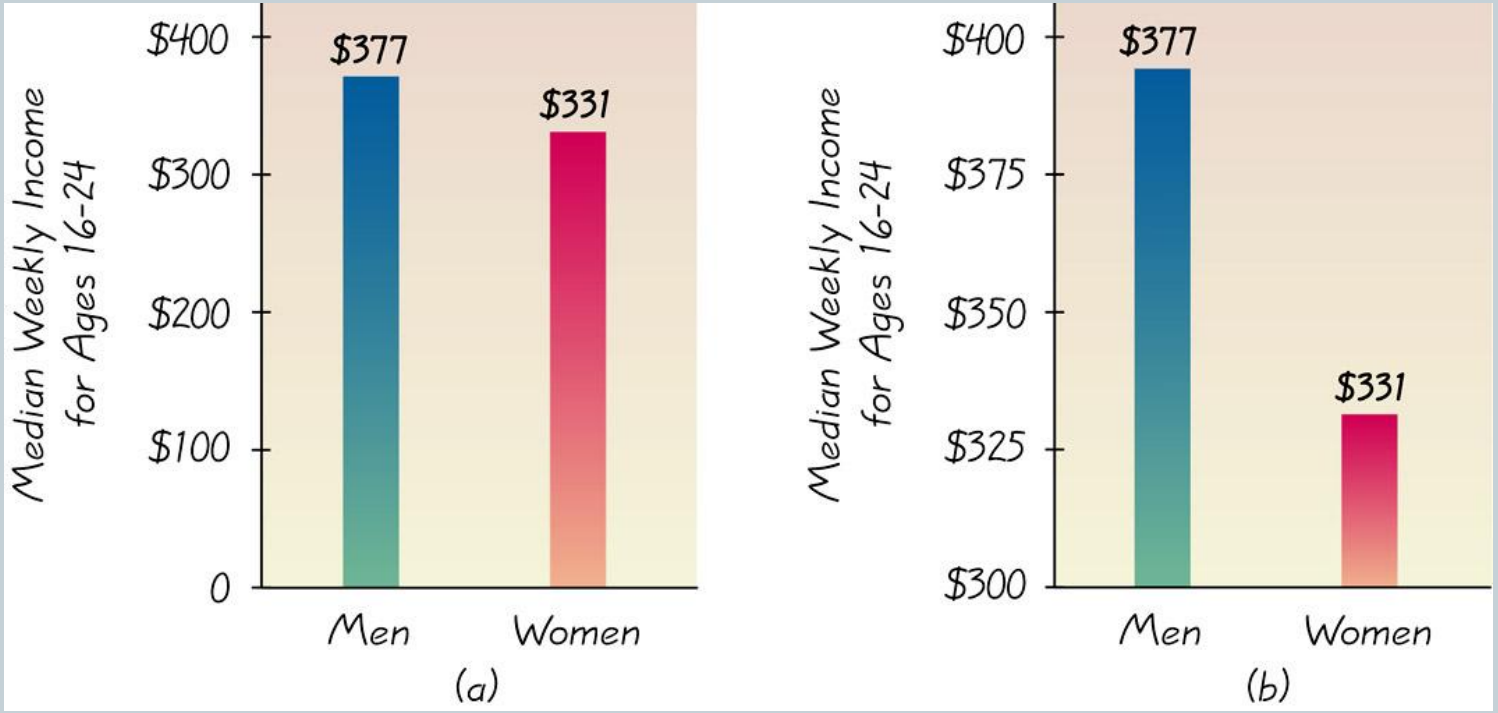
- Voluntary response sample (or self-selected survey)
- The respondents themselves decide whether they want to be included.
- In this case, valid conclusions can be made only about the specific group of people who agreed to participate.

Misuse, Small samples

- Sample→Population
  - The should be representative, for example it should represent the Molecular Biology & Genetics students.
  - Picking two students from the body of
    - 37 departments and 37*30 students in each class and 37*30*4 classes (freshman, sophomore, junior, senior)= about 4500 students
  - is no good. Two students do not represent 4500 students.
- Picking two NBA players and measuring how tall they are.
- You might pick the short point guard players and seriously misinterpret height of the basketball players.

Misuse

- Misleading graphs (look at the Y axis starting point)



Look at the numbers in the graph instead of being mislead by its general shape.

Introduction to Experimental Design

NOT JUST FOR MASTERS STUDENTS ANYMORE ☺

Data collection

- If sample data are not collected in an appropriate way, the data may be so completely useless that no amount of statistical tutoring can salvage them.
- Randomness typically plays a critical role in determining which data to collect.

Observational Study

- Observing and measuring specific characteristics without attempting to modify the subjects being studied.
  - Animal Behavior
  - Human Behavior
  - Plant growth with no treatment

Observational Study

- An observational study draws inferences about the possible effect of a treatment on subjects
- There is no treated group versus a control group.
- This is in contrast with experiments, such as randomized controlled trials, where each subject is assigned to a treated group or a control group.
- Try to reduce the biases of observational studies by quantifying the outcomes
  - Bone density, the amount of some cell or substance in the blood, physical strength or endurance is measured instead of asking questions to the subjects or having a professional observer's opinion.

Observational Study, Some examples

- Animal Behavior
- Human Behavior
- Investigating a link between abortion and breast cancer
- Plant growth with no treatment
- Effects of smoking ban on public health
- Effects of calcium consumption on gender of babies.
  - Experiment on mice, observation on humans
    - K. Celik, S. Serbest, S. Vurur, **A. Pala** and K. Daglioglu. 2003. Experiments to Investigate the Factors that affect the Rate of Sex Constitution. P. Journal of Nutrition. 2(4): 238-241.

Experimental Study

- Experiment: Apply some treatment and then observe its effects on the subjects
- Controlled experiments
  - Compares the results from treatment group to control group, which are practically the same except the independent variable tested
    - Gender, breed, drug tested, age, UV treatment etc.
  - Example: in a drug trial, the group receiving the drug is the experimental group (treatment group); and the one receiving the placebo or regular treatment is the control group.
- Field experiments
  - Outcomes are observed in a natural setting.
  - For this reason, field experiments may have higher external validity than laboratory experiments.
  - Field experiments suffer from contamination:
    - Experimental conditions can be controlled with more precision and certainty in the lab.

Experimental Study

- Cross Sectional Study:
  - Data are observed, measured, and collected at one point in time.
- Prospective (or Longitudinal or Cohort) Study:
  - Data are collected in the future from groups (called cohorts) sharing common factors.
  - A cohort is a group of people who share a common characteristic or experience within a defined period
    - are born, are exposed to a drug or vaccine or pollutant, or undergo a certain medical procedure.
- Cohort studies can either be conducted prospectively, or retrospectively from archived records

Experimental Study

- Retrospective (or Case Control) Study:
  - Data are collected from the past by going back in time.
- Confounding:
  - Occurs in an experiment when the experimenter is not able to distinguish between the effects of different factors

Try to plan the experiment so confounding does not occur.

Some definitions in Experimental Design

- Blind studies
  - The subject does not know he or she is receiving a treatment or placebo
- Double Blind studies
  - Pepsi example

Some definitions in Experimental Design

- Blocks
  - Groups of subjects with similar characteristics
- Completely Randomized Experimental Design
  - Subjects are put into blocks through a process of random selection
- Rigorously Controlled Design
  - Subjects are very carefully chosen

Replication and Parallel

- Replication
  - The replication is needed to get an idea of how variable the results are.
  - Repetition of an experiment when there are enough subjects to recognize the differences in different treatments.
    - We want to generalize our results.
- Parallel
  - The measurements are made on the same stock.

Replication and Parallel

- Replication: samples are drawn from different glasses of milk (experimental units, or replicates) or from different cows or from different udders of the same cow.
- Parallel: samples are from the same glass of milk, or from the same udder of the cow.

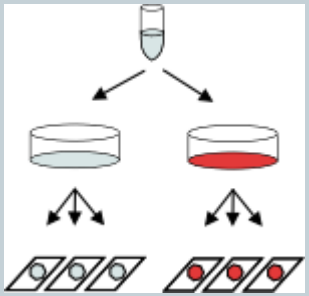
Replication and Parallel

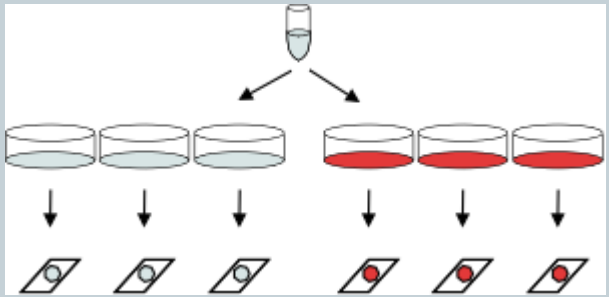
- I have a single mouse, and I make several measures on that mouse
  - multivariate, but not statistics because you have only one mouse.
- I have a litter of mice, and the results can generalized to this parents' offspring.
  - Mice from other parent might be different for this trait
  - Mice from different strains and breeds may be different
- I have different litters from the same inbred strain, so I can generalize to this strain.

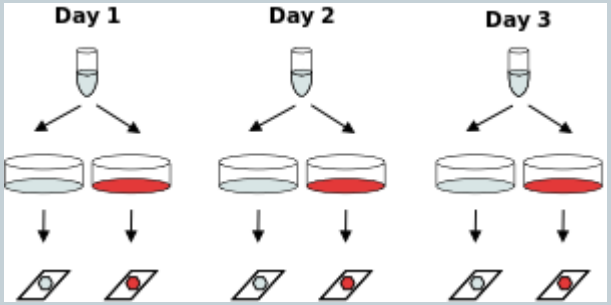
Replication and Parallel

- If 20 blood samples are drawn from 20 people and 10 of them are control, 10 are treatment, n=20
- Here a cell culture or bacteria stock:

No good | better but no cigar | not perfect but fine







- Cells in two culture dishes from the same stock and processed identically do not become fully independent just because a bit of plastic has been placed between them (no cigar or under par).
- With cell lines, there are no biological replicates, and so Design 3 is the best that can be done. The ideal design would have biological replicates (cells from multiple people or animals), and in this case the experiment need only be done once.

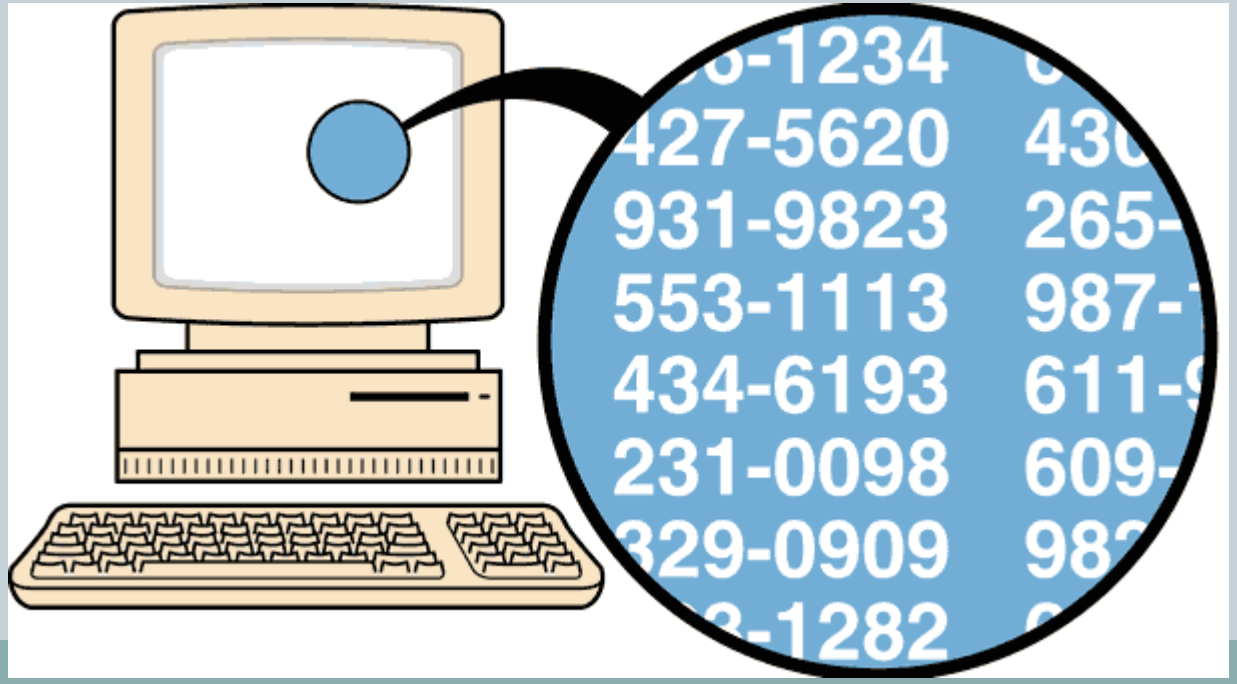
Sample Size

- Sample Size
  - Use a sample size that is large enough to see the true nature of any effects and obtain that sample using an appropriate method, such as one based on randomness
  - We will come back to this.

Sampling with or without replacement

Random Sampling

- Random Sample
- A sample drawn in such a way that each element of the population has a chance of being selected is called a *random sample*.



Random Sampling with replacement

- A sample may be selected with or without replacement.
- Sampling **with replacement:**
  - Each time we select an element from the population, we put it back in the population before we select the next element.
  - The population contains the same number of items each time a selection is made.
  - We may select the same item more than once in such a sample.
- Consider a box that contains 25 marbles of different colors.
- **Rolling a die:** every roll has the same six possible outcomes.
- Bootstrapping for discrete variables.

Random Sampling without replacement

- Sampling **without replacement occurs when the selected element is not replaced in the** population.
- In this case, each time we select an item, the size of the population is reduced by one element.
  - Thus, we cannot select the same item more than once in this type of sampling.
- Most of the time, samples taken in statistics are without replacement.
  - Voters vote only once ☺

Systematic Sampling

- Select some starting point and then select every Kth element in the population

Convenience Sampling

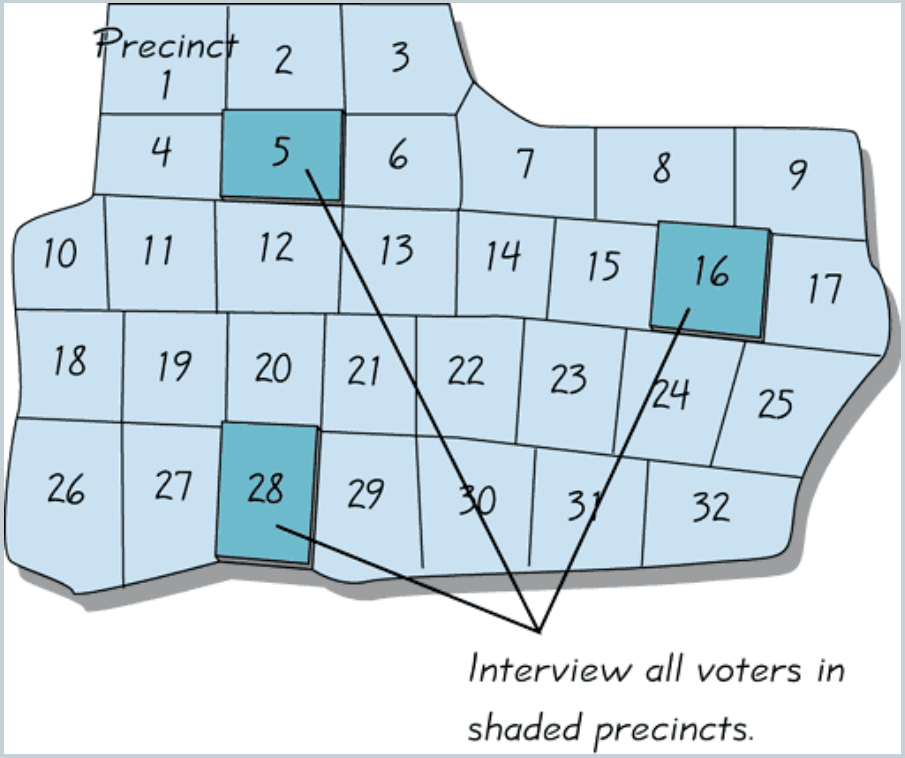
- Use results that are easy to get

Stratified Sampling

- Subdivide the population into at least two different subgroups that share the same characteristics, then draw a sample from each subgroup (or stratum)

Cluster Sampling

- Divide the population into sections (or clusters)
  - randomly select some of those clusters;
  - choose all members from selected clusters: family selection



Methods of Sampling without replacement

- Random
- Systematic
- Convenience
- Stratified
- Cluster

Sampling error

- Sampling Error
  - The difference between a sample result and the true population result; such an error results from chance sample fluctuations
- Nonsampling Error
  - Sample data that are incorrectly collected, recorded, or analyzed (such as by selecting a biased sample, using a defective instrument, or copying the data incorrectly)vv

Homework

- Design an experiment and be sure to include replications.

Misuse, Misinterpretation: correlation

- The correlation phenomena could be caused by a third, previously unconsidered phenomenon, called a confounding variable.
- There is no way to immediately infer the existence of a causal relationship between the two variables.
  - Correlation does not imply causation.
- A correlation between age and height in children is clear
- A correlation between mood and health in people is not
  - Does improved mood lead to improved health, or does good health lead to good mood, or both? Or does some other factor underlie both? A correlation can be taken as evidence for a possible causal relationship, but cannot indicate what the causal relationship, if any, might be.