Statistical Inference

Lecturer: Abdol-Hossein Vahabie

Spring Semester 1401-1402

Writing Assignment I | Deadline 1402/01/09

1 Sampling Strategies

In the following examples classify the sampling method as Census, Simple Random Sample, Cluster, Stratified, or Multistage. Please explain why you chose the classification you did. Additionally, do you think this is the correct approach? If not, what approach would you suggest?

1. There is a myth that science students possess a stronger memory than students in other academic fields. Sara wants to test this hypothesis. Therefore, she designs a creativity test and selects thirty students from each department of the university of Tehran as participants. (3 pts)

2. Suppose you are conducting a study on cognitive abilities among undergraduate students in a large city. First, you randomly select ten universities from a list of all universities in that city. Then, within each selected university, you randomly select fifty students. (3 pts)

3. We want to investigate the effects of a new mindfulness-based intervention on the stress levels of employees in a large organization. we could randomly select five departments from the organization, and then invite all employees in those selected departments to participate in the study. We would administer the mindfulness intervention to the selected employees and compare their stress levels to employees in other departments that did not receive the intervention. (3 pts)

4. A team of researchers distributes flyers on the street for partaking in a memory test. (3 pts)

5. Conducting a survey among the entire student of the University of Tehran to inquire about their predisposition towards participating in a cognitive assessment. (3 pts)

2 Probability Question

The distribution of hospital patients with brain injuries, categorized by age range and the specific area of brain injury within the four main lobes of the brain, is as follows:

	Damaged cerebral lobe				
Age Range	Insular/Limbic	Frontal	Temporal	Parietal	Occipital
00-02 yrs	0.1%	0.7%	1%	3.2%	8.8%
03-12 yrs	0.5%	1.2%	1.5%	1.7%	2%
13-25 yrs	0.3%	6.7%	2.9%	0.5%	0.3%
26+ yrs	4%	36%	28.2%	0.3%	0.1%

1. Given the information that the patient brought to the hospital unconscious, has a damaged working memory, what is the probability that they are in 13-25 age range? (3 pts)

2. What is the probability that a patient will have no dysfunction in the temporal lobe of the brain? (3 pts)

3. What is the probability that an adolescent patient will have Frontotemporal damage? Is this probability feasible to calculate? (3 pts)

4. What is the probability that a patient will have no injury in any of the cortical lobes of the brain? (3 pts)

3 Confounding Variables

Define the term "confounding variable" and assess whether there is one present in the following observations: (3pts)

1. On January 29, 2023, it was observed that the sales of A.S. Roma Football Club's uniforms were several times higher than normal days. On the previous day, a picture of a celebrity wearing Roma's uniform was published on social networks, and many users attributed the rise in sales to this celebrity. (5pts)

2. On October 27, 2022, Elon Musk was forced to buy Twitter after a long conflict with the former owners. After taking over the company, Elon Musk fired top Twitter execs, including the CEO, as well as many talented engineers and employees. However, on November 17, Musk claimed that Twitter had hit an all-time high in usage. Do you believe this happened because of his "great" management skills? (5pts)

3. Some studies suggest that the suicide rate is lower during wartime. Do you suspect that there is a confounding factor in these studies? (5pts)

4. Some studies have found that overweight or obese patients with heart disease have a lower risk of mortality compared to those who are of normal weight. This paradoxical relationship has also been observed in other conditions such as chronic kidney disease, chronic obstructive pulmonary disease (COPD), and some types of cancer. (5pts)

4 The Marshmallow Experiment

Walter Mischel of Stanford University set out to study whether deferred gratification can be an indicator of future success. During the Marshmallow Experiment conducted in 1972, children between the ages of four and six were brought into a room and presented with a marshmallow on a table in front of them. Prior to leaving the room, the researcher informed each child that they would be given a second marshmallow if they refrained from eating the first one for fifteen minutes until the researcher returned. The researcher recorded the amount of time each child was able to resist eating the marshmallow and also noted whether this ability was associated with success in adulthood. Out of the six hundred children tested, a small proportion immediately consumed the marshmallow, while one-third delayed gratification long enough to receive the second marshmallow. In follow-up studies, Mischel discovered that those who exhibited delayed gratification were notably more competent and achieved higher SAT scores compared to their peers.

1. Define the the explanatory variable and the response variable in this experiment? (3 pts)

2. Is this an experimental or an observational study? Why? (3 pts)

3. Do you see any confounding factors in this study? explain your reasons. (3 pts)

4. How can you improve this experiments? pay attention to every element of the study i.e. sampling strategy, confounding variables, etc. (3 pts)

5 Misinformation Analysis

Select a piece of news that contains statistical information and critically analyze it. Look for sources of bias, statistical mistakes, fallacies, misinformation, misleading graphs, and other common deficiencies. Please provide your news source link. (Attach the news link from the news agency e.g. Kayhan website). (8pts)

6 Cognitive Abilities (R)

In a research from Educational Psychology journal [1], cognitive abilities was classified into five categories: memory ability (MA), representational ability (RA), information processing ability (IPA), logical reasoning ability (LRA), and thinking conversion ability (TCA). The study analyzed how these five abilities impacted academic achievement and provided a dataset, which can be accessed [2].

1. Prior to creating histograms for each of the five cognitive abilities, could you make an educated guess regarding the distribution of the data? (6 pts)

2. Create a histogram to visualize the distribution of each attribute and analyze its skewness and modality in your discussion. (6 pts)

3. Create a pie chart to illustrate the frequency of each grade, ensuring that each category is represented by a unique color and percentage value. Additionally, include a legend for the chart to provide clarity. (6 pts)

4. Construct a boxplot for the Total Score (ts) of each gender (sex) group, and use your visual analysis to determine if there are any gender-based differences in academic performance? Later, in this course, you will learn to statistically prove your conclusion. (6 pts)

5. The authors of the paper hypothesized that cognitive abilities positively influence academic achievement. Use visual analysis to explore this idea. Later, in this course, you will learn to statistically prove your conclusion. (6 pts)

Required Document

[1] Shi, Y., Qu, S. (2022). Analysis of the effect of cognitive ability on academic achievement: Moderating role of self-monitoring. In Frontiers in Psychology (Vol. 13). Frontiers Media SA. https://doi.org/10.3389/fpsyg.2022.996504

[2] Download dataset

General Rules

Please upload a file in ZIP format (~~not RAR~~) to the elearn platform(`https://elearn5.ut.ac.ir/course/view.php?id=14838`).

You are allowed a total grace period of 3 days to submit late assignments for all of your exercises.

It is prohibited to use ~~handwritten material~~ and only material produced through typing in the HW template is permissible.

Utilizing a LaTeXto compose the report will grant an additional 5 points.

Please use only built-in R packages for plotting, and ensure that each plot has a title and appropriate legends, if needed and do not use non-built-in packages such as ~~ggplot2~~

Deadline

Wednesday 23:59. 1402/01/09.

Contact Information

Please direct your questions regarding Homework 1 only to the teaching assistants, Mohammad Javad Ranjbar and Sara Rostami, through the course mail (statistical.inference.ut@gmail.com). Use ”HW1” as the subject line.

Good Luck