

Homework 1

Francisco Brady

Exercise 1.2: Sinusitis and antibiotics, Part I.

Researchers studying the effect of antibiotic treatment for acute sinusitis compared to symptomatic treatments randomly assigned 166 adults diagnosed with acute sinusitis to one of two groups: treatment or control. Study participants received either a 10-day course of amoxicillin (an antibiotic) or a placebo similar in appearance and taste. The placebo consisted of symptomatic treatments such as acetaminophen, nasal decongestants, etc. At the end of the 10-day period, patients were asked if they experienced improvement in symptoms. The distribution of responses is summarized below.

		Self-reported improvement in symptoms		
Group	Treatment	Yes	No	Total
	Control	66	19	85
		65	16	81
		131	35	166

(a) What percent of patients in the treatment group experienced improvement in symptoms?

66/85

[1] 0.7764706

(b) What percent experienced improvement in symptoms in the control group?

65/81

[1] 0.8024691

(c) In which group did a higher percentage of patients experience improvement in symptoms?

The control group.

(d) Your findings so far might suggest a real difference in effectiveness of antibiotic and placebo treatments for improving symptoms of sinusitis. However, this is not the only possible conclusion that can be drawn based on your findings so far. What is one other possible explanation for the observed difference between the percentages of patients in the antibiotic and placebo treatment groups that experience improvement in symptoms of sinusitis?

The observed difference between the percentages could be due to sample selection, random variation, or other variables that are not controlled for in the analysis.

Exercise 1.3: Air pollution and birth outcomes, study components.

Researchers collected data to examine the relationship between air pollutants and preterm births in Southern California. During the study air pollution levels were measured by air quality monitoring stations. Specifically, levels of carbon monoxide were recorded in parts per million, nitrogen dioxide and ozone in parts per hundred million, and coarse particulate matter (PM10) in $\mu g/m^3$. Length of gestation data were collected on 143,196

births between the years 1989 and 1993, and air pollution exposure during gestation was calculated for each birth. The analysis suggested that increased ambient PM10 and, to a lesser degree, CO concentrations may be associated with the occurrence of preterm births.

- (a) Identify the main research question of the study.
Does exposure to air pollutants before birth have an impact on the incidence of preterm births in Southern California.
- (b) Who are the subjects in this study, and how many are included?
Subjects of the study are people who gave birth in Southern California and their births between 1989 and 1993.
- (c) What are the variables in the study? Identify each variable as numerical or categorical. If numerical, state whether the variable is discrete or continuous. If categorical, state whether the variable is ordinal.
 - Level of carbon monoxide, numerical, continuous
 - Level of nitrogen dioxide, numerical, continuous
 - Level of ozone, numerical, continuous
 - PMP10, numerical, continuous
 - Length of gestation, numerical, discrete (days?)
 - Air pollution exposure during gestation, categorical

1.4 Buteyko method, study components.

The Buteyko method is a shallow breathing technique developed by Konstantin Buteyko, a Russian doctor, in 1952. Anecdotal evidence suggests that the Buteyko method can reduce asthma symptoms and improve quality of life. In a scientific study to determine the effectiveness of this method, researchers recruited 600 asthma patients aged 18-69 who relied on medication for asthma treatment. These patients were randomly split into two research groups: one practiced the Buteyko method and the other did not. Patients were scored on quality of life, activity, asthma symptoms, and medication reduction on a scale from 0 to 10. On average, the participants in the Buteyko group experienced a significant reduction in asthma symptoms and an improvement in quality of life.

- (a) Identify the main research question of the study.
Can the Buteyko method reduce symptoms of asthma and improve quality of life?
- (b) Who are the subjects in this study, and how many are included?
600 asthma patients, aged 18 to 69.
- (c) What are the variables in the study? Identify each variable as numerical or categorical. If numerical, state whether the variable is discrete or continuous. If categorical, state whether the variable is ordinal.
 - Quality of life, numerical, discrete
 - activity, numerical, discrete
 - asthma symptoms, numerical, discrete
 - medication reduction, numerical, discrete

1.7: Migraine and acupuncture, Part II.

Exercise 1.1 introduced a study exploring whether acupuncture had any effect on migraines. Researchers conducted a randomized controlled study where patients were randomly assigned to one of two groups: treatment or control. The patients in the treatment group received acupuncture that was specifically designed to treat migraines. The patients in the control group received placebo acupuncture (needle insertion at non-acupoint locations). 24 hours after patients received acupuncture, they were asked if they were pain free. What are the explanatory and response variables in this study?

- Explanatory variable: acupuncture treatment

- Response variable: Response to pain free question

1.12: UN Votes.

The visualization below shows voting patterns in the United States, Canada, and Mexico in the United Nations General Assembly on a variety of issues. Specifically, for a given year between 1946 and 2015, it displays the percentage of roll calls in which the country voted yes for each issue. This visualization was constructed based on a dataset where each observation is a country/year pair.

- List the variables used in creating this visualization.
 - yes votes for each country
 - topic of each vote
 - year of vote
- Indicate whether each variable in the study is numerical or categorical. If numerical, identify as continuous or discrete. If categorical, indicate if the variable is ordinal.
 - year, numerical, discrete
 - topic/issue of vote, categorical
 - yes/no vote, categorical

1.15: Buteyko method, scope of inference.

Exercise 1.4 introduces a study on using the Buteyko shallow breathing technique to reduce asthma symptoms and improve quality of life. As part of this study 600 asthma patients aged 18-69 who relied on medication for asthma treatment were recruited and randomly assigned to two groups: one practiced the Buteyko method and the other did not. Those in the Buteyko group experienced, on average, a significant reduction in asthma symptoms and an improvement in quality of life.

- Identify the population of interest and the sample in this study. Asthma patients, the sample is 600 asthma patients between the ages of 18 and 69.
- Comment on whether or not the results of the study can be generalized to the population, and if the findings of the study can be used to establish causal relationships. The results of the study cannot be generalized to the population, because the study was only conducted on asthma patients. It's not clear whether the results would extend to non-asthma patients. In order to establish a causal relationship, the study would have to control for other variables.

1.17: Relaxing after work.

The General Social Survey asked the question, "After an average work day, about how many hours do you have to relax or pursue activities that you enjoy?" to a random sample of 1,155 Americans. The average relaxing time was found to be 1.65 hours. Determine which of the following is an observation, a variable, a sample statistic (value calculated based on the observed sample), or a population parameter.

- An American in the sample.
Observation
- Number of hours spent relaxing after an average work day.
Variable
- 1.65.
Sample statistic
- Average number of hours all Americans spend relaxing after an average work day.
Population parameter