Section 1.2 – Observational Studies versus Designed Experiments

Distinguish between an Observational Study and an Experiment

Basics of Collecting Data

Statistical methods are driven by the data that we collect. We typically obtain data from two distinct sources: *observational studies* and *experiment*.

Example

Researchers Joachim Schüz and associates wanted "to investigate cancer risk among Danish cellular phone users who were followed for up to 21 years." To do so, they kept track of 420,095 people whose first cellular telephone subscription was between 1982 and 1995. In 2002, they recorded the number of people out of the 420,095 people who had a brain tumor and compared the rate of brain tumors in this group to the rate of brain tumors in the general population.

Example

They found no significant difference in the rate of brain tumors between the two groups. The researchers concluded "cellular telephone was not associated with increased risk for brain tumors." (*Source*: Joachim Schüz et al. "Cellular Telephone Use and Cancer Risk: Update of a Nationwide Danish Cohort," *Journal of the National Cancer Institute* 98(23): 1707-1713, 2006)

Researchers Joseph L. Roti and associates examined "whether chronic exposure to radio frequency (RF) radiation at two common cell phone signals–835.62 megahertz, a frequency used by analogue cell phones, and 847.74 megahertz, a frequency used by digital cell phones–caused brain tumors in rats. The rats in group 1 were exposed to the analogue cell phone frequency; the rats in group 2 were exposed to the digital frequency; the rats in group 3 served as controls and received no radiation. The exposure was done for 4 hours a day, 5 days a week for 2 years. The rats in all three groups were treated the same, except for the RF exposure.

After 505 days of exposure, the researchers reported the following after analyzing the data. "We found no statistically significant increases in any tumor type, including brain, liver, lung or kidney, compared to the control group." (*Source*: M. La Regina, E. Moros, W. Pickard, W. Straube, J. L. Roti Roti. "The Effect of Chronic Exposure to 835.62 MHz FMCW or 847.7 MHz CDMA on the incidence of Spontaneous Tumors in Rats." Bioelectromagnetic Society Conference, June 25, 2002.)

In both studies, the goal of the research was to determine if radio frequencies from cell phones increase the risk of contracting brain tumors. Whether or not brain cancer was contracted is the ***response variable***. The level of cell phone usage is the ***explanatory variable***.
In research, we wish to determine how varying the amount of an ***explanatory variable*** affects the value of a ***response variable***.

Definitions

An ***observational study*** measures the value of the response variable without attempting to influence the value of either the response or explanatory variables. That is, in an observational study, the researcher observes the behavior of the individuals in the study without trying to influence the outcome of the study.

Example: The Literary Digest poll in which respondents were asked who they would vote for in the presidential election is an observational study. The subjects were asked for their choices, but they were not given any type of treatment.

If a researcher assigns the individuals in a study to a certain group, intentionally changes the value of the explanatory variable, and then records the value of the response variable for each group, the researcher is conducting a ***designed experiment***.

Experiment apply some treatment and then observe its effects on the subjects; (subjects in experiments are called experimental units)

Example: In the largest public health experiment ever conducted, 200,745 children were given a treatment consisting of the Salk vaccine, while 201,229 other children were given a placebo. The Salk vaccine injections constitute a treatment that modified the subjects, so this is an example of an experiment.

Example

Researchers wanted to determine the long-term benefits of the influenza vaccine on seniors aged 65 years and older. The researchers looked at records of over 36,000 seniors for 10 years. The seniors were divided into two groups. Group 1 were seniors who chose to get a flu vaccination shot, and group 2 were seniors who chose not to get a flu vaccination shot. After observing the seniors for 10 years, it was determined that seniors who get flu shots are 27% less likely to be hospitalized for pneumonia or influenza and 48% less likely to die from pneumonia or influenza. (Source: Kristin L. Nichol, MD, MPH, MBA, James D. Nordin, MD, MPH, David B. Nelson, PhD, John P. Mullooly, PhD, Eelko Hak, PhD. "Effectiveness of Influenza Vaccine in the Community-Dwelling Elderly," New England Journal of Medicine 357:1373–1381, 2007)

Based on the results of this study, would you recommend that all seniors go out and get a flu shot?
The study may have flaws! Namely, *confounding.*

Definitions

Confounding in a study occurs when the effects of two or more explanatory variables are not separated. Therefore, any relation that may exist between an explanatory variable and the response variable may be due to some other variable or variables not accounted for in the study.

A ***lurking variable*** is an explanatory variable that was not considered in a study, but that affect the value of the response variable in the study. In addition, lurking variables are typically related to any explanatory variables considered in the study.

- ✓ Some lurking variables in the influenza study: age, health status, or mobility of the senior
- ✓ Even after accounting for potential lurking variables, the authors of the study concluded that getting an influenza shot is *associated* with a lower risk of being hospitalized or dying from influenza.
- ✓ Observational studies do not allow a researcher to claim causation, only association.

Explain the Various Types of Observational Studies

Cross-sectional Studies Observational studies that collect information about individuals at a specific point in time, or over a very short period of time.

Case-control Studies These studies are ***retrospective***, meaning that they require individuals to look back in time or require the researcher to look at existing records. In case-control studies, individuals who have certain characteristics are matched with those that do not.

Cohort Studies A cohort study first identifies a group of individuals to participate in the study (the cohort). The cohort is then observed over a long period of time. Over this time period, characteristics about the individuals are recorded. Because the data is collected over time, cohort studies are ***prospective***.

Example

Determine whether each of the following studies depict an observational study or an experiment. If the researchers conducted an observational study, determine the type of the observational study.

a) Researchers wanted to assess the long-term psychological effects on children evacuated during World War II. They obtained a sample of 169 former evacuees and a control group of 43 people who were children during the war but were not evacuated. The subjects' mental states were evaluated using questionnaires. It was determined that the psychological well being of the individuals was adversely affected by evacuation. (Source: Foster D, Davies S, and Steele H (2003) The evacuation of British children during World War II: a preliminary investigation into the long-term psychological effects. Aging & Mental Health (7)5.)

Observational study; Case-control

b) Xylitol has proven effective in preventing dental caries (cavities) when included in food or gum. A total of 75 Peruvian children were given milk with and without xylitol and were asked to evaluate the taste of each. Overall, the children preferred the milk flavored with xylitol. (Source: Castillo JL, et al (2005) Children's acceptance of milk with xylitol or sorbitol for dental caries prevention. BMC Oral Health (5)6.)

Designed experiment

c) A total of 974 homeless women in the Los Angeles area were surveyed to determine their level of satisfaction with the healthcare provided by shelter clinics versus the healthcare provided by government clinics. The women reported greater quality satisfaction with the shelter and outreach clinics compared to the government clinics. (*Source*: Swanson KA, Andersen R, Gelberg L (2003) Patient satisfaction for homeless women. Journal of Women's Health (12)7.)

Observational study; Cross-sectional

d) The Cancer Prevention Study II (CPS-II) is funded and conducted by the American Cancer Society. Its goal is to examine the relationship among environmental and lifestyle factors on cancer cases by tracking approximately 1.2 million men and women. Study participants completed an initial study questionnaire in 1982 providing information on a range of lifestyle factors such as diet, alcohol and tobacco use, occupation, medical history, and family cancer history. These data have been examined extensively in relation to cancer mortality. Vital status of study participants is updated biennially.

Cause of death has been documented for over 98% of all deaths that have occurred. Mortality follow-up of the CPS-II participants is complete through 2002 and is expected to continue for many years. (*Source*: American Cancer Society)

Observational study; cohort

Definition

A ***census*** is a list of all individuals in a population along with certain characteristics of each individual.

Exercises *Section* 1.2 – Observational Studies vs Designed Experiments

1. Researchers wanted to know if there is a link between proximity to high-tension wires and the rate of leukemia in children. To conduct the study, researchers compared the rate of leukemia for children who lived within ½ mile of high-tension wires to the rate of leukemia for children who did not live within ½ mile of high-tension wires. Determine whether the study depicts an observational study or an experiment.

2. Rats with cancer are divided into two groups. One group receives 5 milligrams (mg) of a medication that is thought to fight cancer, and the other receives 10 mg. After 2 years, the spread of the cancer is measured. Determine whether the study depicts an observational study or an experiment.

3. Seventh-grade students are randomly divided into two groups. One group is taught math using traditional techniques; the other is taught math using a reform method. After 1 year, each group is given an achievement test to compare proficiency. Determine whether the study depicts an observational study or an experiment.

4. A poll is conducted in which 500 people are asked whom they plan to vote for in the upcoming election. Determine whether the study depicts an observational study or an experiment.

5. A survey is conducted asking 400 people. "Do you prefer Coke or Pepsi?" Determine whether the study depicts an observational study or an experiment.

6. A Gallup poll surveyed 1018 adults by telephone, and 22% of them reported that they smoked cigarettes within the past year. Determine whether the description corresponds to an observation study or an experiment.

7. In a morally and criminally wrong study, 399 black men with syphilis were not given a treatment that could have cured them. The intent was to learn about the effects of syphilis on black men, The subjects were initially treated with small amounts of bismuth, neoarsphenamine, and mercury, but those treatments were replaced with aspirin. Determine whether the description corresponds to an observation study or an experiment.

8. While shopping, 200 people are asked to perform a taste test in which they drink from two randomly placed, unmarked cups. They are then asked which drink they prefer. Determine whether the description corresponds to an observation study or an experiment.

9. Conservation agents netted 250 large-mouth bass in a lake and determined how many were carrying parasites. Determine whether the description corresponds to an observation study or an experiment.

10. Researchers wanted to determine if there was an association between the level of happiness of an individual and their risk of heart disease. The researchers studied 1739 people over the course of 10 years. During this 10-year period, they interviewed the individuals and asked questions about their daily lives and the hassles they face. In addition, hypothetical scenarios were presented to determine how each individual would handle the situation. These interviews were videotaped and studied to

assess the emotions of the individuals. The researchers also determined which individuals in the study experienced any type of heart disease over the 10-year period. After their analysis, the researchers concluded that the happy individuals were less likely to experience heart disease.

a) What type of observational study is this? Explain.
b) What is the response variable?
c) What is the explanatory variable?
d) In the report, the researchers stated that "the research team also hasn't ruled out that a common factor like genetics could be causing both the emotions and the heart disease." Use the language introduced on this section to explain what this sentence means.

11. Researchers wanted to determine if there was an association between daily coffee consumption and the occurrence of skin cancer. The researchers looked at 93,676 women enrolled in the Women's Health Initiative Observation Study and asked them to report their coffee-drinking habits. The researchers also determined which of the women had nonmelanoma skin cancer. After their analysis, the researchers concluded that consumption of six or more cups of caffeinated coffee per day was associated with a reduction in nonmelanoma skin cancer

a) What type of observational study is this? Explain.
b) What is the response variable?
c) What is the explanatory variable?
d) In their report, the researchers stated that "After adjusting for various demographic and lifestyle variables, daily consumption of six or more cups was associated with a 30% reduced prevalence of nonmelanoma skin cancer." Why was it important to adjust for these variables?

12. Researcher Penny Gordon-Larson and her associate wanted to determine whether young couples who marry or cohabitate are more likely to gain weight than those who stay single. The researchers followed 8000 men and women for 7 years as they matured from teens to young adults. When the study began, none of the participants were married or living with a romantic partner. By the end of the study, 14% of the participants were married and 16% were living with a romantic partner. The researchers found that married or cohabiting women gained, on average, 9 pounds more than single.

a) Why is this an observation study? What type of observational study is this?
b) What is the response variable in the study?
c) What is the explanatory variable?
d) Identify some potential lurking variables in this study.
e) Can we conclude that getting married or cohabiting causes one to gain weight? Explain.