MOD 4 TYPES OF STATISTICAL STUDIES

Learning Goals

- Determine whether a question is about a population or about cause-and-effect.
- Determine if a study is an experiment or an observational study.
- Explain how the study design impacts the types of conclusions that can be drawn.
- Suppose a History instructor allows students to select one of the two topics below for a term paper they are required to write. Further suppose about half the class selects the first topic and half selects the second topic. Which group of students do you think would be more successful when writing their term paper? Explain.
 - a) Write a term paper on the history of the world.b) Write a term paper about the problems Martin Luther King needed
 - problems Martin Luther King needed to overcome in order to lead the civil rights movement in the United States.
- 2) Suppose we are researchers tasked with investigating the relationship between smoking and lung capacity. We begin by writing research questions that will guide our investigation. For each of the following research questions, decide whether the question would make it easier or harder for us to collect and analyze data to answer the question. Then rank all four questions from best (rank 1) to worst (rank 4). Explain your rankings, i.e. explain why the better ranked question(s) would make it easier to collect and analyze data, and why the worst ranked question(s) would make it more difficult to collect and analyze data.
 - a) How does smoking affect lung capacity of women?
 - b) On average, do women who smoke have less lung capacity than women who do not smoke?
 - c) Does smoking cause reduced lung capacity in women?
 - d) Are women who smoke more likely to ruin their lung capacity?

- 3) Copy your top two ranked research questions below, and then label each one as a question about a *population* or a question about *cause and effect*.
- 4) For each of your top two questions what are the groups being compared?
- 5) For each of your top two questions what is the *variable* we are trying to measure?
- 6) Two statistical studies are described below (we'll improve on each of the studies in subsequent activities, but for now, let's explore what we have).

a)	a)
b)	b)
c)	c)
 Find 100 women age 20 who do not smoke. Randomly assign 50 of the 100 women to the smoking treatment and the other 50 to the no-smoking treatment. Those in the smoking group smoke a pack a day for 10 years, and those in the control group remain smoke-free for 10 years. Measure and record lung capacity for each of the 100 women. Analyze, interpret, and draw conclusions from the data. 	 Find 100 women age 30 such that 50 of them have been smoking a pack a day for 10 years and 50 of them have been smoke-free for 10 years. Measure lung capacity for each of the 100 women. Analyze, interpret, and draw conclusions form the data.

a) In the first box at the top of each column, label the statistical study as an observational study or an experiment (I know ... we haven't defined these terms yet ... but give it your best shot). What features of the study made you label it an observational study or an experiment?

- b) In the first box at the top of each column on the previous page indicate whether or not the study would provide convincing evidence about cause and effect. In the space below, explain why you labeled each study as *provides convincing evidence about cause and effect* or *does NOT provide convincing evidence about cause and effect*.
- c) Match your top two research questions to the corresponding study in the table on the previous page (write the research question in the first box at the top of the appropriate column).
- 7) In each of the statistical studies listed on the previous page, what are the variables (note – the variables are the same in both studies)? Which is the *explanatory variable* and which is the *response variable*? Explain.
- 8) Let's take a little sidetrack ... Suppose you're part of a group of 10 friends that get together on Sundays to play basketball. Two people in your group are really good players and two people in your group are pretty bad at the game (but they've got heart and they are the glue that keeps the group together). Everyone else is fairly average.
 - a) Now suppose that one team gets both of the really good players and each team gets one bad player. What do you suspect will likely happen when the two teams play each other? Explain and try to use the word "bias" in your explanation.
 - b) Now suppose that one team gets both of the bad players and each team gets one really good player. What do you suspect will likely happen when the two teams play each other? Explain and try to use the word "bias" in your explanation.
 - c) Now suppose that each team gets one really good player and each team gets one bad player. Is the outcome of the game biased toward either team winning? Explain why or why not.

9) Suppose you are part of a research group tasked with investigating the question, "<u>Are</u> people who smoke more likely to develop lung cancer?"

Before developing your research plan, your group learns about a gene that makes the smell and taste of tobacco irresistible, AND also makes people more likely to develop lung cancer regardless of whether they smoke or not. In other words the presence of the imaginary gene makes smoking very desirable and makes lung cancer more likely whether the person smokes or not. Unfortunately the test to determine which subjects in your study have the gene and which do not is astronomically expensive. So your research group cannot afford to test each subject for this gene.

a) Suppose your research group conducts an observational study. You recruit 300 people who have been smoking one pack of cigarettes per day for the past 10 years and 300 people who have not smoked for the past 10 year? Then you compare the lung cancer rate between the smoker and non-smoker groups. Would the gene bias the outcome? If so, how? If not, why not?

b) Suppose your research group will conduct an experiment to investigate the question, "<u>Does smoking increase the risk of lung cancer?</u>" What can you do mitigate the effect of the gene that makes people who have the gene want to smoke and simultaneously makes them more likely to develop lung cancer? Be prepared to share your conclusions with the entire class.