***Coursera – Intro to Probability***

*Week 1 Quizzes*

1. Which of the following classifications of variable types is false?

* Customer satisfaction: very unsatisfied, unsatisfied, satisfied, very satisfied → ordinal categorical
* **Population of each state in the US → continuous numerical = FALSE**
* *If variable is numerical, it is then continuous or discrete based on whether or not it can take on an infinite # of values or only non-negative whole numbers, respectively.*
* *If variable is categorical, determine if it is ordinal based on whether or not the levels have a natural ordering.*
* *Counted data are discrete numerical variables since they can’t take on non-whole values.*
* Student height → continuous numerical
* Whether a student has previously taken a statistics course → categorical

1. If subjects are randomly assigned to treatments, conclusions can be generalized to the population.

* True
* **False**
* *ideal experiment = random sampling AND assignment = causal conclusions that CAN be generalized to a whole population*
* *most experiments = volunteer subjects = NO random sampling, but random assignment = causal conclusions, but cannot be generalized.*
* *typical observational study = NO random assignment, but uses random sampling = correlation statements but generalized to the population at large.*

*UN-ideal observational study = does NOT use random assignment OR random sampling + can only be used to make correlational statements that are NOT generalizable*

1. As part of a stats project, Andrea would like to collect data on household size in her city. To do so, she asks each person in her stats class for the size of their household, + reports that her sample is a simple random sample. However, it’s not a simple random sample. Which of the following is the best reasoning for why this is not a random sample that is appropriate for this research question?

* Andrea did not block for any variables that might influence the response.
* **Andrea did not use any randomization; she took a convenience sample.**
* Andrea did not use a stratified sample.

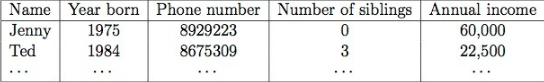
1. Which of the following is not one of the four principles of experimental design?

* replicate
* randomize
* **stratify**
* *CONTROL any possible confounders*
* *RANDOMIZE into treatment + control groups*
* *REPLICATE by using a sufficiently large sample or repeating the experiment*
* *BLOCK any variables that might influence the response.*

1. **Stratified sampling** allows for controlling for possible confounders in the sampling stage, while blocking allows for controlling for such variables during random assignment.

* **True**
* False
* *Stratifying + blocking BOTH allow for controlling for potential confounders, but at different stages of the study design.*
* *Stratify when sampling (divide population into strata + sample from w/in each stratum)*
* *Block during random assignment (divide sample into blocks + randomly assign from w/in each block to treatment groups).*

1. Consider table below describing a data set of individuals registered to volunteer at a public school. Which of the choices below lists categorical variables?



* name and number of siblings
* number of siblings and year born
* **phone number and name**
* annual income and phone number

1. A study is designed to test the effect of type of light on exam performance of students. 180 students are randomly assigned to 3 classrooms: 1 dimly lit, another w/ yellow lighting, and a 3rd w/ white fluorescent lighting, + given the same exam. Which of the following correctly ID’s the variables used in the study as explanatory and response?

* explanatory: dimly lit, yellow, white fluorescent

response: exam performance

* **explanatory: type of light (categorical with 3 levels)**

**response: exam performance**

* explanatory: exam performance

response: dimly lit, yellow, white fluorescent

* explanatory: exam performance

response: type of light (categorical with 3 levels)

1. In a study published in 2011 in The Proceedings of the National Academy of Sciences, researchers randomly assigned 120 elderly men + women who volunteered to be a part of this study (average age mid-60s) to 1 of 2 exercise groups. 1 group walked around a track 3 times a week; the other did a variety of less aerobic exercises, including yoga + resistance training w/ bands. After a year, brain scans showed that among the walkers, the hippocampus (part of brain responsible for forming memories) had increased in volume by about 2% on average; in the others, it had declined by about 1.4%. Which of the following is false?

* The explanatory variable is the type of exercise, and the response variable is the change in volume of the hippocampus.
* **The results of this study can be generalized to all elderly 🡪** *not random sampling = not generalized*
* A causal link between walking and expansion of the hippocampus can be inferred based on these results 🡪 *random assignment = causal*

1. An extraneous variable related to the explanatory + response variables that prevents us from deducing causal relationships based on observational studies is called a \_\_\_\_\_\_\_.

* **Confounding variable**

1. As part of a stats project, Andrea would like to collect data on household size in her city. To do so, she asks each person in her stats class for the size of their household, + reports that her sample is a simple random sample. However, this is not a simple random sample. Which of the following is the best reasoning for why this is not a random sample that is appropriate for this research question?

* **In this investigation of household size, each household represents a case. Andrea incorrectly sampled individuals instead of households.**
* Andrea asked everybody in her class instead of asking her classmates to volunteer
* Andrea did not use a random number table to randomize the order in which she collected the students’ responses, so the sample cannot be random.

1. In an experiment, what purpose does blocking serve?

* Increase sample size.
* Obtain a random sample.
* **Control for variables that might influence the response.**
* Prevent skewed results.

1. Which of the following is one of the four principles of experimental design?

* **randomize**
* stratify
* cluster