

Problem 1

The effects of exposure to lead on the psychological and neurological well-being of children were studied by Landrigan et al. (1975). Complete raw data for this study are in the data set lead.csv on Blackboard. The data in lead.csv describe a group of children who lived in one of three areas (area) near a lead smelter in El Paso, Texas. An exposed group of 46 children were identified who had blood-lead levels ≥ 40 $\mu\text{g/ml}$ in 1972 (or for a few children in 1973). This group is defined by the variable GROUP = 1. A control group of 78 children were also identified who had blood-lead levels < 40 $\mu\text{g/ml}$ in both 1972 and 1973. This group is defined by the variable GROUP = 2.

1. How many individuals are in the data set?
124 individuals.
2. How many variables are in this data set?
40 variables.
3. Can you tell if any of the variables are categorical (i.e. qualitative)? Identify specific ones.
In this data frame, such as the first column called "ID" is a categorical variable because ID is fixed, each row has its own ID respectively. We cannot see the same ID in this data frame. Also, although ID is a number, it cannot be calculated so it is not a qualitative variable.

Two important variables that were studied were (1) MAXFT = the number of finger-wrist taps in the dominant hand (a measure of neurological function) and (2) IQF = the Wechsler full-scale IQ score. You will explore the relationship of lead exposure to one of these two outcome variables.

4. Is this an observational study or a randomized experiment? Explain why.
Is an observational study. Because I did not see that any variable is randomly sampled in this data frame. This dataset just like "recording the data with each ID" so we can know each ID has different characteristics.

5. How many individuals have MAXFT scores measured? How many have IQF scores measured?

99 individuals have MAXFT scores measured, 124 individuals have IQF scores measured.

6. Pick one of MAXFT or IQF of interest to you. We are primarily interested in comparing the distribution of the outcome of interest (MAXFT or IQF) for the two different groups of children (GROUPS 1 and 2, those children with elevated blood-lead levels > 40 $\mu\text{g/ml}$ and those with lower levels, < 40 $\mu\text{g/ml}$, respectively.)

- a. What are the mean and median of the outcome of interest (MAXFT or IQF) for each GROUP?

Mean: Group 1 of MAXFT is 47.42857

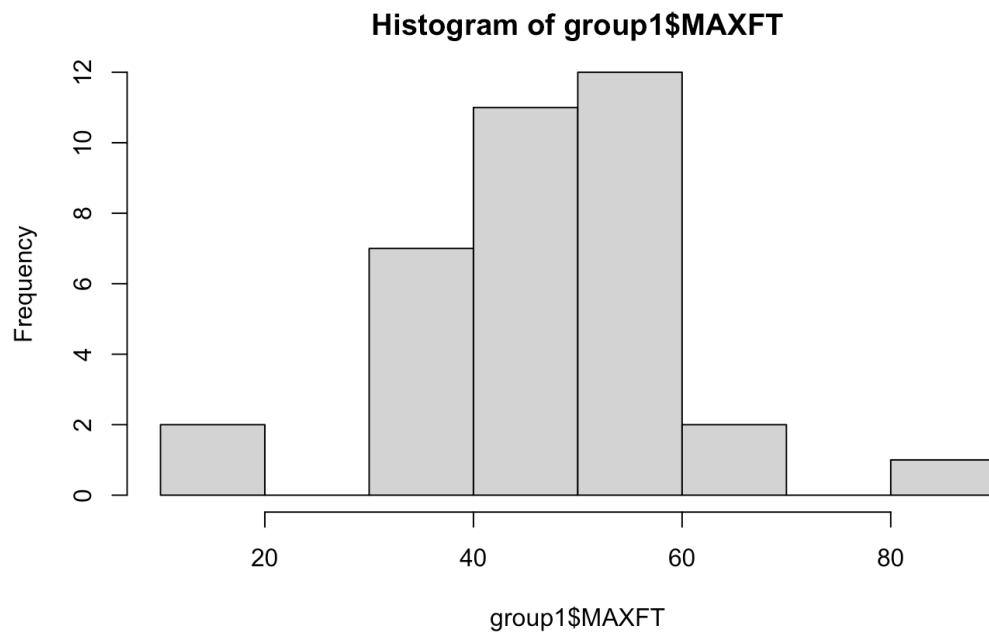
Median: Group 1 of MAXFT is 48

Mean: Group 2 of MAXFT is 54.4375

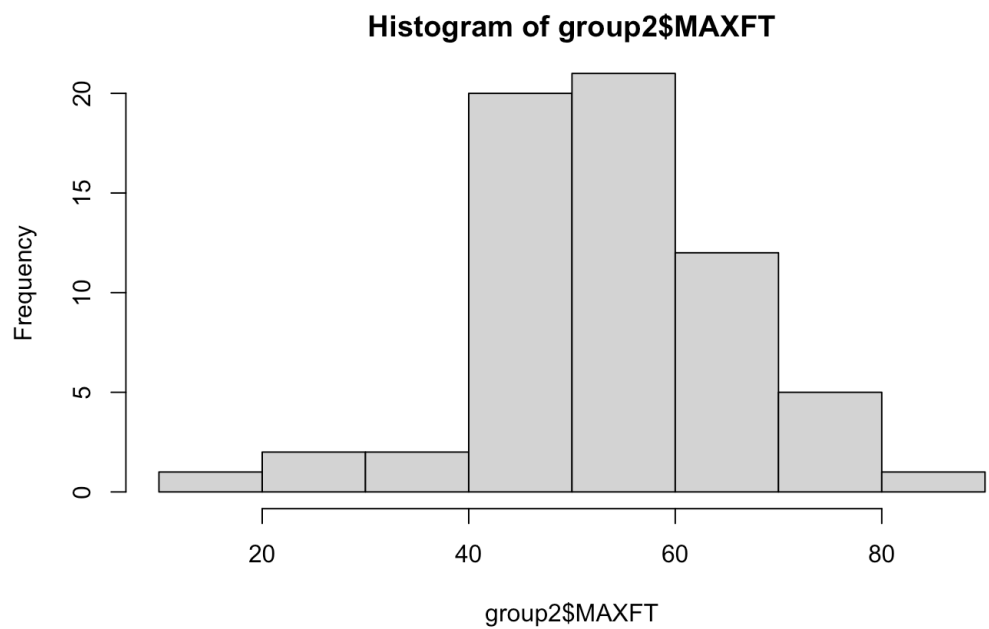
Median: Group 2 of MAXFT is 53.5

- b. Describe the shape of the distribution (i.e. histogram) of the outcome for each GROUP.

Group1: children with elevated blood-lead levels $\geq 40 \mu\text{g/ml}$
The concentration of blood-lead levels has not interfered with the flexibility of their finger.

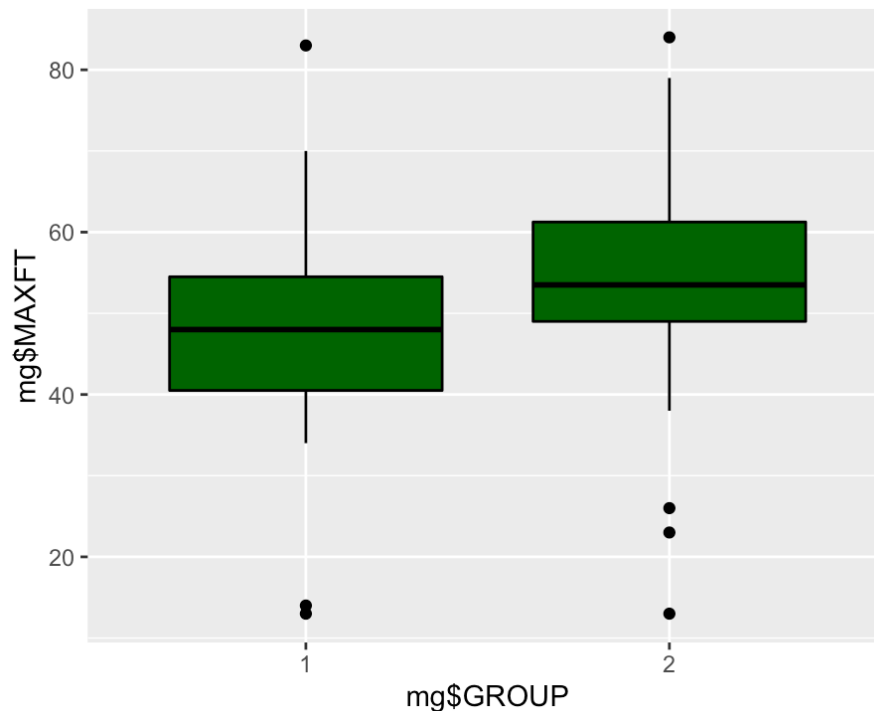


Group2: children with elevated blood-lead levels $< 40 \mu\text{g/ml}$
The concentration of blood-lead levels has not interfered with the flexibility of their finger.



c. What information can we get from the Boxplot of the outcome for each GROUP?

Group2 has more wide range, but the median between group 1 and 2 are similar.



d. Based on these summaries, what is your assessment of the differences between the two groups of children on the outcome of interest? Discuss the role of randomization in this study.

It is not a randomization for this analysis, because we use its originally observed data to distinguish group 1 and group 2. Although there is no significant difference between 1 and 2, group 2 has more data (total 64) that we can be referred to, compared to group 1 it only has 35 people.

References:

- Rosner B, (2005) Fundamentals of Biostatistics, Thomson.
- Landrigan PJ, Whitworth RH, Baloh RW, Staehling NW, Barthel WF, Rosenblum BF (1975) Neuropsychological dysfunction in children with chronic low-level lead absorption. Lancet, 1, 708-715.

Problem 2

Re-visit the study design for Case Study 1.1.2 Sex Discrimination in Employment from Chapter 1 of the textbook or our class notes. Briefly contrast the case study design with that of the study described in the New York Times article "Bias Persists for Women of Science, a Study Finds" and in the manuscript, Science Faculty's Subtle Gender Biases Favor Male Students (both are given in Blackboard under the Homework tab.) Briefly describe the overall goals of each study and give the strengths and weaknesses of each. Which study do you find has more compelling evidence for the hypotheses of interest? **(Dr. Malloy: One of the main differences between the two studies is the that Science Faculty study involved a randomized experiment.)**

We can know that there is some sex discrimination between these two articles. In the second article, the article has taken the randomization of double-blind study (n = 127) so they have

enough evidence to support their points, also we can see the mean and sd with each group. (Male target student and Female target student.) In fig.2, we saw this graphic shows the massive difference between male's salaries and female's salaries, and although females and males have the same ability, people have preferred to choose males and pay them a higher salary. Compared to the first article called *Bias Persists for Women of Science, a Study Finds*, they still gave the result of investigations, such as the average starting salary with different sex, but if they can insert some plots to explain the association between both in the article that would be more clear for the readers.

For both articles, did the faculty and professor have a different opinion if we asked the same question on a second time? Because sometimes we have some bias for answering questions or being influenced by the person who is beside you. If so, this investigation would be a little bit different.

Problem 3

Please take the brief and anonymous Introductory Survey located in Blackboard under the Homework tab.

I have done.

Problem 4

Please send me an email and tell me at least one thing about yourself. This can be anything - you get to decide what you would like to share with me. It could be about your hopes and wishes for this class, any concerns you have about taking your courses online, a fun fact about yourself I wouldn't otherwise learn...whatever you want me to know about you. You can share whatever you want to share!

Also, feel free to ask me something in this email, whether related to the course or not. My hope is this will be the beginning of a dialogue and you become comfortable communicating with me in a variety of formats. I will respond to your email – and I try to respond to all email within 24 hours but because I have nearly 50 students this semester it may take me a little longer this time to get back to you!

I have done.