Statistical Thinking for Forensic Practitioners

Quiz on Part 3: Data Collection

1 Randomization of subjects

Randomization is useful for ensuring that any differences between subjects of a study are evened out across treatment groups. We will explore how different randomization techniques can be used to control for differences among subjects. Click on the following link to access a Randomization applet.

Once you open the applet, you should see on the right a set of 16 blue and 8 red cards with names and heights of hypothetical participants in a study. The blue and red cards correspond to male and female participants, respectively. Below the 24 cards are the phrases "Group 1" and "Group 2." We will be analyzing how a measured response differs between the two groups based on how we randomize the assignment to these groups. To the left of these cards you should see a drop-down arrow next to the word "sex". Click this drop-down arrow and select "height" instead. We will be treating the particpants' heights as our measured response. Also click the "Show Graphs" button to change the data summary to a plot.

The purpose of using randomization to assign subjects to groups is to render the two groups similar/homogeneous, on average. We will quantify how similar the two treatment groups by considering the difference between the two groups' mean height values. With this in mind, complete the following problems.

1. Calculate the difference between the two groups' mean height values based on this initial group assignment. Take note of this original assignment, in particular that all female participants are assigned to Group 2.

   The difference is about 4.67. All female participants are assigned to Group 2.

2. Click the "Randomize" button once. An animation will play showing the 24 cards being randomly "dealt" into two groups. Record the difference in the sample means based on this new assignment.

   Answers will vary.

3. How does this compare to the difference calculated in question 1? Do you believe this would happen again if we were to repeat the random assignment? Explain.

   The original assignment was sort of a "worst-case" scenario in that all female participants were assigned to the same group. As females are shorter than males on average, we would expect that the difference calculated in question 1 is larger than the difference of effectively any other assignment.

Click the "Animate" checkbox to turn off the card dealing animation. Change the number next to "Replications" from 1 to 10 (and make sure "Animate" is still unchecked) and click the "Randomize" button 10 more times. As you do, notice the behavior of the histogram southwest of the "Randomize" button.

4. Around what value does this histogram appear to be centered? Considering the purpose of using randomization to assign groups, explain why this value makes sense.

   The distribution seems centered on 0. This is the precisely the "evening out of differences" phenomenon for which randomization is used.

5. Calculate the total number of possible assignments of these 24 participants into two equal-sized groups. (Hint: see slide 15 of the Part 3 lecture slides. You can use this website to calculate the answer)

   Choosing 12 participants to go into Group 1 exactly determines which participants go into Group 2. Thus, there are "24 choose 12" total combinations. $\binom{24}{12} = 2704156$ **assignments**

6. Click the "Reset" button near the bottom of the page to reset the app. We will now consider what happens when we control the group assignment for sex. Considering that our variable of interest is height, why might it be useful to control the group assignment for sex?

   Females tend to be shorter than males on average, so we may not want to treat them interchangeably when making group assignments if our goal is to make the groups as "homogeneous" as possible.

To control for sex, we will enforce the restriction that every assignment have the same number of males and females. In particular, we will separately assign the 16 males and 8 females into the two groups so that each group consists of 8 randomly selected males and 4 randomly selected females. This is commonly referred to as a "Randomized Block" design where the "Blocks" correspond to sex in this example. Click the "Randomized Block" button on the left side of the page. Click the "Randomize" button a few times to observe how the cards are now assigned into the two groups.

7. Now that we have controlled for sex, how do you expect the standard deviations of the differences between the groups' height values to compare to those calculated under the "Completely Randomized" design used in questions 2-4? Explain.

   If we are enforcing the groups to be more homogeneous using this Randomized Block design, then we would expect the differences to be less variable than in the Completely Randomized design case.

8. To test your answer from question 6, we will simulate many possible group assignments. Are there more or fewer possible assignments using the Randomized Block design than the Completely Randomized design? Explain. (Note: you don't need to actually calculate the number of possible assignments under the Randomized Block design)

   Given that we are restricting the types of assignment we can make, there are most certainly fewer possible assignments. This is sufficient to answer the question. However, to figure this out mathematically, we can apply similar logic as used in the solution to problem 5. Selecting 8 out of the 16 males and 4 out of the 8 females to go into Group 1 exactly determines which participants go into Group 2. Thus, for every assignment of males, there are "8 choose 4" possible assignments of females. Since there are "16 choose 8" possible assignments of males, the total number of possible assignments is then "16 choose 8" multiplied by "8 choose 4."

   $\binom{16}{8} \cdot \binom{8}{4} = 900900$ **assignments**

Change the number next to "Replications" from 10 to 1000 (and make sure "Animate" is still unchecked). Click the "Randomize" button. While we were simulating assignments using the Randomized Block design, the applet was also secretly simulating assignments from the Complete Randomized design behind-the-scenes. To see the group difference histograms under both of these designs, click the "Both" button on the left side of the screen. The Completely Randomized histogram is shown in light gray behind the Randomized Block histogram.

9. Considering the shapes of these two histograms (and the standard deviations computed), does anything surprise you about the results? Explain. (Note: this answer may simply be "No" if the results are similar to what you expected in question 6.)

   Answers will vary. The spread of the Randomized Block histogram is smaller than that of the Completely Randomized histogram. This was to be expected due to the enforced homogeneity of the assigned groups under the Randomized Block design.

2 Representative sampling and the scope of inference

Consider a 2003 Research Report in the *Journal of Epidemial Community Health* by Shenassa et al. called "Lethality of firearms relative to other suicide methods: a population based study" (Source). The objectives of the research were "(1) [t]o quantify lethality of firearms relative to other suicide methods, (2) to quantify the extent to which suicide mortality may be reduced by limiting access to firearms." The data upon which the analysis was based were gathered during the period of 1990 to 1997 from mortality files of the Chicago Department of Public Health and from hospital discharge data in the state of Illinois.

10. Is this analysis an observational or a randomized controlled study? (Note: refer to slide 33 of the Part 3 lecture slides.)

    An observational study

11. Given your answer to question 10, comment on the second goal of the analysis. To what degree can we use this study to draw a causal relationship between accessibility to firearms and the prevalence of suicide mortality (i.e., does access to firearms cause an increase in fatal suicides attempts)?

    It is not possible to draw a causal relationship between between accessibility to firearms and the prevalence of suicide mortality because assignment of treatments (suicide by firearm vs. not) was not randomized (or "assigned," for that matter). This makes it difficult to draw any strong conclusions about the authors' second goal.

12. Is it possible (legally, ethically, etc.) to carry out a controlled study to accomplish the goals of this analysis? Explain.

    It would not be possible to carry out a controlled study. It would certainly violate ethical and legal standards and likely not even be physically possible to carry out.

13. What do you believe to be the population of interest for this study? (Note: there may be multiple correct answers here. This is intended to stimulate your thinking.)

    It's not entirely clear from the paper what the population of interest is. Assuming the authors are interested in generalizing their findings to the largest population possible (they are performing "a population based study," after all), they might consider their population to be all individuals who desire to commit suicide. Alternatively, the population need not be a population of people, but rather could be a population of events. One could argue that the (widest) population of the interest is all suicide attempts (fatal and non-fatal). This question was mainly intended to spur your thinking to transition into the next question.

14. An anti-bullying advocacy group from Los Angeles, CA wants to use this study's data for a study involving suicide rates among Californian teens in the age of social media. Comment on aspects of the data that might affect their applicability to this study.

    The fact that the original data were collected in Illinois from 1990 to 1997 means that they are likely not particularly applicable to a more modern study based on California. For example, firearm laws differ by state and change over time, so perhaps the availability of more lethal firearms differs between the time/place in which these studies took place. Additionally, we know nothing about the age demographics of the Illinois dataset whereas the anti-bullying advocacy group is focused specifically on teenagers. This might also affect the applicability of the original data.