14.1: Refresh to Prepare

We are heading into a whole new territory now! In the first part of this book, we looked at how to describe data with the measures of central tendency (mean, median, and mode), the measure of variability (standard deviation), and graphing. Then in the second part of this book, we looked at using inferential statistics to compare means with t-tests and ANOVAs. That's a short sentence to describe all that you learned in that second section!

? Exercise 14.1.1

What kind of variables can we calculate a mean from: Quantitative or Qualitative?

Answer

Means can only be calculated for quantitative variables.

Now, we are leaving means behind, and focusing on *relationships* between variables.

But before we do that, let's make sure that we remember the difference between qualitative and quantitative variables.

? Exercise 14.1.2

What type of variable is each of the following: Quantitative or Qualitative?

1. Major in college
2. Age
3. Gender
4. Test scores (in points)
5. Ounces of vodka

Answer

1. Major is qualitative (named categories or groups)
2. Age is quantitative (number)
3. Gender is qualitative (named categories or groups)
4. Test scores (in points) is quantitative (number)
5. Ounces of vodka is quantitative (number)

In this section of the textbook, we are going to focus on health. If we wanted to compare the nutritional value of different food items at different fast food restaurants, we could run a variety of statistical analyses.

As a reminder, a t-test has one quantitative variable (DV) and one qualitative variable with two levels (IV).

✓ Example 14.1.1

If we wanted to compare the average calories for cheeseburgers from different two different fast food restaurants:

1. What is the quantitative DV? In other words, what is being measured?
2. What is the IV and its two levels? The IV levels are the qualitative groups.

Solution

1. The DV would be the average calories for the cheeseburgers. Calories are numeric, so they are quantitative.
2. The IV is the two different restaurants.

An ANOVA also has one quantitative variable and one (or more) qualitative variables.

Example 14.1.2

If we wanted to compare the average calories for three or more types of food (cheeseburger, chicken sandwich, salad) from one fast food restaurant:

1. What is the quantitative DV? In other words, what is being measured?
2. What is the IV and its levels? The IV levels are the qualitative groups.

Solution

1. The DV would be the average calories for the cheeseburgers. Calories are numeric, so they are quantitative.
2. The IV is the different types of food, with the levels being cheeseburger, chicken sandwich, or salad.

(Or more...)

Exercise 14.1.3

If we wanted to compare the average calories for three or more types of food (cheeseburger, chicken sandwich, salad) from two different fast food restaurants:

1. What is the quantitative DV? In other words, what is being measured?
2. What is one of the IVs and its levels? The IV levels are the qualitative groups.
3. What is one of the IVs and its levels? The IV levels are the qualitative groups.
4. What kind of factorial design would that be? 2x2? 2x3? 2x2x2? Something else?

Answer

1. The DV would be the average calories for the cheeseburgers. Calories are numeric, so they are quantitative.
2. One IV is type of food, and the levels are cheeseburger, chicken sandwich, or salad. These are all qualitatively different foods, which are different categories.
3. Another IV is the restaurant, with two different restaurants being the two levels.
4. 2x3: Restaurant (2 levels) by Food Type (Cheeseburger, Chicken Sandwich, or Salad)

In this chapter, we will analyze scenarios in which we have only two quantitative variables. Although we could look at the means of each different variable, it doesn't really make sense to compare differences in means of two different variables. For example, let's say that our two variables were cheeseburger calories and cheeseburger price. We could find the average calories of all of the cheeseburger was 350 calories, and the average prices was $4.49. Without doing any statistical analyses, I can say that 350 is higher than 4.49, so comparing means doesn't give us any useful information when we have two quantitative variables. Remember way back in the section on IVs and DVs when we talked about IVs being predictors and DVs being outcomes? The IV is the predictor that we think is what we causes the outcome, which is the DV.

This is still true with two quantitative variables. We can predict that the IV of calories predicts the outcome (DV) of cost. We can use inferential statistics to test is if prices go up when calories go up. Thus, the research question changes from "Are these means from different populations?" to something more like, "Do these numbers vary together?" We follow the Null Hypothesis Significance Testing process that we've been doing:

1. Step 1: State the Hypotheses
2. Step 2: Find the Critical Values
3. Step 3: Compute the Test Statistic
4. Step 4: Make the Decision, then report the results!

Now that you've been refreshed, let's move on to our next statistical analysis: Pearson's Correlation