SIPPS Coding Workshop: advanced track
01: Refresher of regression and moderation (Frequentist)

Hannah & Camille

June 8, 2021

Before we begin

- ▶ make sure you have R/RStudio downloaded on your machines! (if you don't, check out this tutorial: https://cu-psych-computing.github.io/cu-psych-comp-tutorial/tutorials/r-core/1-programming/lessonpart0/)
- ▶ please let us know if we're ever moving too fast! (or too slow)
- ▶ any and all questions are welcome, at any time
- ▶ recordings of this workshop, plus annotated notes of the code we'll be going through today, will be available shortly after the session

Plan for today

- what is a linear regression model? (quick review)
- prepping your data for analysis
- using the `lm()` package & interpreting the output
- visualizing your results

What is a linear regression model?

- ▶ used to quantify the relationship(s) between an **outcome variable** and one (or more!) **predictors** (also referred to as **moderators**)

What is a linear regression model?

▶ used to quantify the relationship(s) between an **outcome variable** and one (or more!) **predictors** (also referred to as **moderators**)

▶ for example, let's say you are interested in which variables affect how much you'll enjoy a particular take-out restaurant

What is a linear regression model?

▶ used to quantify the relationship(s) between an **outcome variable** and one (or more!) **predictors** (also referred to as **moderators**)

▶ for example, let's say you are interested in which variables affect how much you'll enjoy a particular take-out restaurant

▶ you could hypothesize that your enjoyment will depend on (at least) two things:

What is a linear regression model?

- used to quantify the relationship(s) between an **outcome variable** and one (or more!) **predictors** (also referred to as **moderators**)

- for example, let's say you are interested in which variables affect how much you'll enjoy a particular take-out restaurant

- you could hypothesize that your enjoyment will depend on (at least) two things:
  - the average Yelp rating

What is a linear regression model?

▶ used to quantify the relationship(s) between an **outcome variable** and one (or more!) **predictors** (also referred to as **moderators**)

▶ for example, let's say you are interested in which variables affect how much you'll enjoy a particular take-out restaurant

▶ you could hypothesize that your enjoyment will depend on (at least) two things:
  ▶ the average Yelp rating
  ▶ whether the restaurant is a pizzeria or a taqueria

What is a linear regression model?

▶ used to quantify the relationship(s) between an **outcome variable** and one (or more!) **predictors** (also referred to as **moderators**)

▶ for example, let's say you are interested in which variables affect how much you'll enjoy a particular take-out restaurant

▶ you could hypothesize that your enjoyment will depend on (at least) two things:
  ▶ the average Yelp rating
  ▶ whether the restaurant is a pizzeria or a taqueria

▶ `enjoyment ~ (yelp rating) + (type of restaurant)`

Types of variables

- regression models are inherently flexible, and allow you to quantify many different kinds of variables & relationships

Types of variables

- regression models are inherently flexible, and allow you to quantify many different kinds of variables & relationships

- in our toy model, for example, we are looking at two different types of moderators/predictors: **continuous** (the yelp rating) & **categorical** (the type of restaurant)

Types of variables

▶ example results:
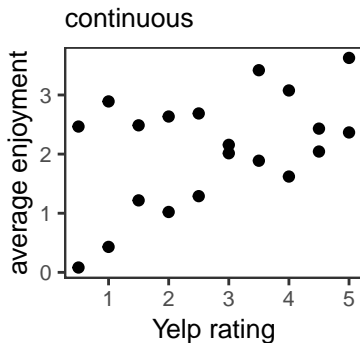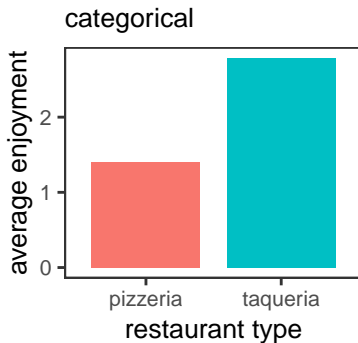
Types of effects

- there are two main "classes" of effects that regression models can reveal

Types of effects

▶ there are two main "classes" of effects that regression models can reveal

▶ **main effects**: the value of one predictor (by itself) has an impact on the outcome variable

Types of effects

- ▶ there are two main "classes" of effects that regression models can reveal

- ▶ **main effects**: the value of one predictor (by itself) has an impact on the outcome variable

- ▶ **interactions**: the impact of one predictor on the outcome variable *depends* on the value of another variable

Types of effects

- there are two main "classes" of effects that regression models can reveal

- **main effects**: the value of one predictor (by itself) has an impact on the outcome variable

- **interactions**: the impact of one predictor on the outcome variable *depends* on the value of another variable

- `enjoyment ~ (yelp rating) * (type of restaurant)`

Types of effects

▶ here we see the interaction: the importance of Yelp rating to overall enjoyment *depends* on the type of restaurant
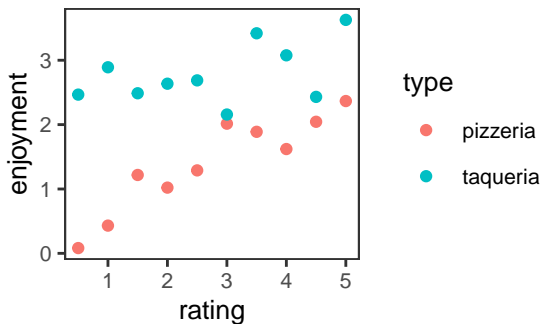
Plan for today

- ~~what is a linear regression model?~~ any questions?
- prepping your data for analysis
- using the `lm()` package & interpreting the output
- visualizing your results