CH. 11 – Analysis of Variance (ANOVA)

PY 221 Statistics & Research Methods I

Dr. Valenti

Outline for Ch. 11 – Analysis of Variance (ANOVA)

1. Overview of when to use ANOVA
2. What is ANOVA (*F*-ratios – aka, "omnibus tests"), conceptually?
   - What question do they help us answer?
   - Null and alternative hypotheses
   - How are *F*-ratios calculated?
3. Comparing subsets of means
   i. Post-hoc tests
   ii. Planned contrasts

REVIEW

- What do ***t*-tests** allow you to do?
  - examine whether the apparent *differences* between two groups on an outcome variable are real differences (vs. produced by chance)
  - examine whether the apparent *changes in people's scores* across two time points are real changes (vs. produced by chance)
  - predict the value of an outcome variable based on the category (group) it came from, when there are two categories (groups)

ANOVA: for when you want to assess differences across several (more than two) means

- What does **ANOVA** allow you to do?
  - examine whether apparent *differences* between three or more groups on an outcome variable are real differences (vs. produced by chance)
  - predict the value of an outcome variable based on the category (group) it came from, when there are three or more categories (groups)

Not true for t-tests

- can use with qualitative/categorical predictors with 3+ categories
- can use with more than one qualitative predictor
- can use for within-subjects designs

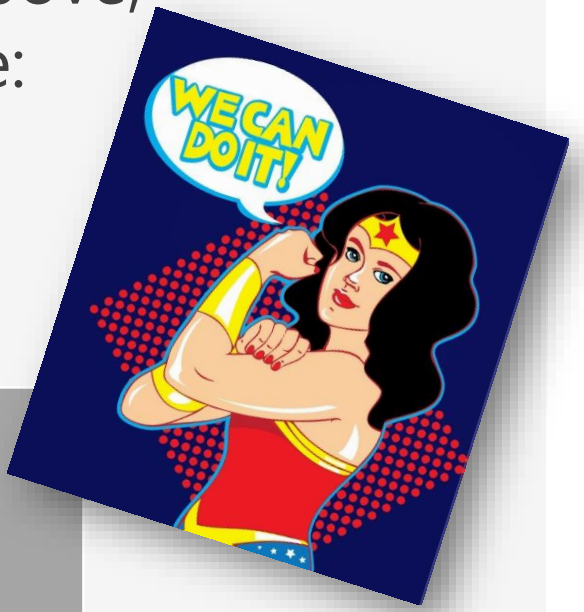
In this course, we won't cover how to run analyses in the latter two situations.

Research questions that would involve ANOVA - examples

1. Does taking **notes by hand**, taking **notes w/a laptop**, or **not taking notes** in class cause better understanding of the material?
2. Does watching **Stephen King's *It***, the **news**, ***Homeland***, or **a tarantula crawling around a tank** produce greater anxiety?
3. Do **vegan diets** or **non-vegan diets** lead to greater muscle strength in college students, and does it depend on whether the student **plays or does not play** a college sport?
    - Here we have *two predictor variables* – diet and whether play sports, while the first two examples have one predictor each.

What causes students to have a "can-do" attitude in statistics class?

- Outcome/dependent variable:
  - A classroom observer will rate each student's attitude during class on a scale ranging from 1-*"can't do"* to 10-"*can-do"*
- Generate a research question with the outcome variable above, that would require ANOVA. The study design must involve:
  - 1 qualitative predictor with 3+ levels *or*
  - 2 qualitative predictors (with 2+ levels each)



think (write) → **pair** → **share**

If we want to compare several means, why don't we just compare pairs of means using *t*-tests?

- E.g., compare anxiety from **Stephen King's *It*** with anxiety from the **news**
- E.g., compare anxiety from ***Homeland*** with anxiety from the **news**
- *Etc.*, until all possible pairs are compared?

- Primary reason
  - **The more individual *t*-tests you run, the greater the chance that your sample data will reveal a significant difference between some of the groups when there is no real difference in the population.**
- i.e., using multiple *t*-tests increases the **Type I** error rate!

Practice your understanding

Identify the research question(s) whose data should be analyzed with ANOVA. (Select as many answers as apply).

1. Is a person's job satisfaction (assume satisfaction is measured on a quantitative scale) related to the amount of time (in minutes) it takes for the person to commute to work?
2. Is year in college (freshman, sophomore, etc.) related to a student's GPA?
3. Is amount of money (in $) in one's bank account related to one's heart rate when hearing how much money a car repair will cost?
4. Is the amount of sleep a person gets each night dependent on the particular day of the week?
5. Is whether or not you identify as a female (yes vs. no) related to cat ownership (yes, own vs. no, don't own)?

 Answer:#2 and #4

Outline for Ch. 11 – Analysis of Variance (ANOVA)

1. Overview of when to use ANOVA
2. What is ANOVA (*F*-ratios – aka, "omnibus tests"), conceptually?
   - What question do they help us answer?
   - Null and alternative hypotheses
   - How are F-ratios calculated?
3. Comparing subsets of means
   i. Post-hoc tests
   ii. Planned contrasts

F-ratio: assessing differences between several means

- Take this research question . . .
  - Does taking **notes by hand**, taking **notes w/a laptop**, or **not taking notes** in class cause better understanding of the material, when understanding is measured using a 7-question quiz?

Q: What does an *F*-ratio tell us?

- ***A:*** The relative accuracy of using the *individual* ***group*** *means* to predict an outcome score, versus using the ***overall*** *outcome mean* to predict an outcome score.

F-ratio: assessing differences between several means

- We learn the relative accuracy of using the *individual group means* to predict an outcome score, versus using the *overall outcome mean* to predict an outcome score.
- **Example**:
  - DV/outcome = after-class level of understanding of material (score on 7-question quiz)
  - IV/predictor = method of note-taking during class
    - **Group 1**: notes by hand // **Group 2**: notes w/laptop // **Group 3**: take no notes at all, just listen

Is knowing which **group** a student is in helpful in predicting how well they'll understand the material at the end of class?

Or does method of note-taking (i.e., group) not tell us anything about a student's likelihood of understanding the material?

DV/Outcome: Student level of understanding of material (score on 7-item quiz at the end of class)

- N = 15 students
- *Overall* (i.e., *"grand"*) *mean* for level of understanding = sum of *all* scores/N = **3.47**

- To accurately predict the score of a new student, should we use the grand mean, i.e., 3.47?
- *Or* would it help to know how the student takes notes (i.e., what group they're in)?

Manipulation (IV) is *method of note-taking*

	Group 1	Group 2	Group 3
	By hand	w/laptop	No notes
	7	3	5
	4	2	2
	5	1	4
	3	1	2
	6	4	3

DV/Outcome: Student level of understanding of material (score on 7-item quiz at the end of class)

- N = 15 students
- *Overall/Grand mean* for level of understanding =
  sum of *all* scores/N = 3.47
- To accurately predict the score of a new student, should we use the grand mean, i.e., 3.47?
- *Or* would it help to know how the student takes notes (i.e., what group they're in)?
- The group means *look* different, but are these *real* differences?
  - *Our F-ratio will help us learn the answer to this question.*

Manipulation (IV) is *method of note-taking*

	Group 1	Group 2	Group 3
	By hand	w/laptop	No notes
	7	3	5
	4	2	2
	5	1	4
	3	1	2
	6	4	3
Group Means	5.00	2.20	3.20