All handouts for this class: **https://tinyurl.com/IST772crowston**

IST772 ANOVA (Week 6)

Pre-class activity:

- Open the file 0. Week6anovaVSttest.xlsx from the handouts area.
- Calculate the value of F in the ANOVA table. Report in the chat on the connection between t and F in the chat.

1

The logic of a null hypothesis test

- Collect a sample of data
- Compute a test statistic
  - Often it's some measure of a difference + a scale
  - E.g., t-test statistic is the difference between two means / pooled sd
- Compute the distribution of the test statistic assuming the null hypothesis is true
  - E.g., t-test statistic for no difference follows a t distribution
- See where the test statistic falls in the distribution
  - If it's extreme and unlikely, then some assumption is likely wrong
  - So, we may reject the null hypothesis

2

A Statistician's Thoughts Turn To. . .

- What are the assumptions of the test: viz the nature of the input data (e.g., parametric tests may require normally distributed inputs). If my input data does not fit the assumptions of the test, what then?
- Frequentist: What is the test statistic, how is it distributed, how are data translated into that test statistic?
- Bayesian: What are the priors, what is modeled for the posterior distribution(s), what parameters does the model track, did the model converge?
- What limitations are there on the interpretation of the results? Does all the evidence point in the same direction? What are the impacts of the research design? What are the impacts of measurement?

3

Very Different Modeling Approaches: Frequentist and Bayesian ANOVA

Frequentist ANOVA

- Null hypothesis: groups are sampled from the same distribution.
- Models a ratio of between groups variance to within groups variance
- Under the null hypothesis, these two variance estimates should be equal
- Under the null hypothesis, ratios of variances are distributed as F
- F is a family of distributions; random deviates generally have a long right tail.

Bayesian ANOVA

- Models means for each group relative to the grand mean
- Models the priors for each group mean as normally distributed around the grand mean.
- Models the priors for each group's sigma from the uniform distribution
- Samples from a posterior distribution of values for each group.

4

Two exercises

ANOVA Calculation in Excel

- Open 1. Week6anovaCalcsInExcel.xlsx
- Observe how the ANOVA test statistic is computed
- Exercise: how much do you need to change the data in group 3 for the test to not be significant?

ANOVA Case Studies

- Open 2. Week 6 ANOVA Case Studies.Rmd
- Exercise: In groups, work through the 2 cases
- Use the poll to report when your group is done

5

Breakout 1 – Freq and Bayesian ANOVA

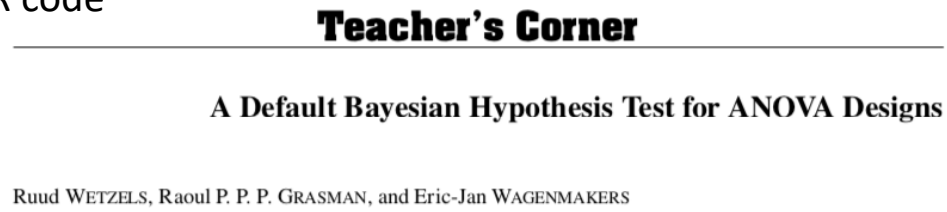
- Open 3. Week 6 Breakout.Rmd
- Insurance data set from Brett Lantz
- Compare medical charges by region
- Run frequentist ANOVA with aov() and Bayesian ANOVA with anovaBF() – run posterior samples for Bayesian model
- Share your code on https://codeshare.io/aJDyRX

All handouts for this class: **https://tinyurl.com/IST772crowston**

6

Paper of the Week – Wetzels et al., 2012

- A Default Bayesian Hypothesis Test for ANOVA Designs
- The details of how Bayesian methods can be applied to ANOVA
- Provides technical details for the specification of priors and the calculation of Bayes factors for ANOVA
- Provides R code



Teacher's Corner

A Default Bayesian Hypothesis Test for ANOVA Designs

Ruud WETZELS, Raoul P. P. P. GRASMAN, and Eric-Jan WAGENMAKERS

7

Homework and Midterm Exam

- Midterm exam file is in the handouts area
- Submit your exam no later than 8:30 PM tomorrow (Wednesday)
- The homework for week 6 is based on exercises 1-7 on pages 117 and 118 but with changes as noted in the notebook. For problem 7, in addition to the Bayesian t-test, feel free to run Tukey's HSD or another post-hoc procedure.

8