Fundamentals of Statistics (MIT OCW 18.6501x (Fall 2016)
- Professor Philippe Rigollet ⟶ MIT
- unit 4 - Hypothesis Testing
- unit 6 - Linear Regression
- unit 2: Inference

I. Probability and Linear Algebra Review
- Gaussian Probabilities
- What is statistics?
  - mathematical theory behind statistical models
  - How to lie w/ statistics
- Probability

Lecture Reading Notes (week 5) 04/01/19

①

CHAPTER 15: ANALYSIS OF VARIANCE

INTRO. - statistical method used to test the difference between 2 or more means
→ inferences about the mean can be made by analyzing the variance
→ used to test general rather than specific differences among means
- ANOVA tests the non-specific null hypothesis that all population means are equal, eg

$\mu_{false} = \mu_{fett} = \mu_{miserable} = \mu_{neutral}$

→ Also called the OMNIBUS NULL HYPOTHESIS
- Rejecting the omnibus null hypothesis means that at least one population mean is different from at least one other mean.
- But we don't know which mean is different from which, so Tukey HSD test more specific

(2)

Chapter 15: Analysis of variance (continuation)

I. ANALYSIS OF VARIANCE DESIGNS

FACTORS & LEVELS:

- Factor = synonymous with independent variable
- levels = types of factors being compared
- one-way ANOVA → conducted on a design where there's only one factor

Between & within-subjects factors

- Between subjects → comparisons b/n diff groups of subjects
- when diff subjects are used for the levels of a factor, the factor is called a "Between-subjects" factor
- within subjects factor → when the same subjects are used for the levels of a factor
  - ↳ "repeated measures variables"
- MULTIFACTOR DESIGNS:
  - factorial design = when all combinations of the levels are included

II. One-Factor ANOVA (Between subjects)

- Null hypothesis tested by ANOVA:

$$H_0: \mu_1 = \mu_2 \cdots = \mu_k$$

k = the number of conditions

- ANOVA = method for testing differences among means by analyzing the variance

1. Mean squared error → based on differences among scores within the groups
2. mean squared between → based on differences among the sample means (only estimates σ^2 if the population means are equal)

- if MSB > MSE, then population means unlikely to be equal
- Assumptions of ANOVA:
  1) populations have the same variance
  2) populations are normally distributed
  3) Each value is sampled independently from each other value

Sample sizes:

- n = # of observations in each group
- N = total # of obs.

Computing MSE

- σ^2 is the quantity estimated by MSE and computed as the mean of sample variances

Computing MSB

- variance of the sampling distribution of the mean

$$\sigma_m^2 = \frac{\sigma^2}{n} \text{, so we get } \sigma^2 = n\sigma_m^2$$

Comparing MSE & MSB

- MSB must be > MSE to conclude that means are not equal — but how much larger?
- R Fisher ⟶ F Ratio
- small sample size = more unstable results
- large sample ⟶ MSE & MSB almost the same
- depends on degrees of freedom for $\frac{MSB}{MSE}$ ⟶ df = $k(n-1) = N-k$

$$df_{numerator} = k-1$$
$$df_{denominator} = N-k$$

- F test = one tailed, but test of two-tailed hypothesis

Relationship to the t test

- when there are only 2 groups

$$F(1, dfd) = t^2(df)$$

Sources of variation

- unexplained variance = error
- var due to the condition the subject was in
- ANOVA — partitions variation into sources
  - ↳ sum of squares (SSQ) = variation

$$SSQ_{total} = \sum (X - GM)^2$$

↑ grand mean → mean of all subjects

$$SSQ_{condition} = n \sum (M_1 - GM)^2 + (M_2 - GM)^2$$

$SSQ_{error} = SSQ_{total} - SSQ_{condition}$

$MSB = SSQ_{condition} / dfn$

$MSE = SSQ_{error} / dfn$

dfd → degrees of freedom error

→ ANOVA summary table
Source, df, SSQ, MS, F, P

IV. Multi-factor Between-Subjects Designs

→ Basic concepts & terms
- interaction: when effect of one var differs depending on level of second var
- main effect of an independent variable is the effect of the variable averaging over the levels of the other variables
  - * marginal means
- simple effect: effect of a var at a single level of another variable
  - * "there is an interaction when the simple effects differ

→ Tests of significance

df of main effect = # of levels - 1

$MS = \frac{SSQ}{df}$

$F\ ratio = \frac{MS\ for\ effect}{MSE}$

df of interaction: product of df's vars in the interaction

error df: N - n
(N: obs, n: groups)

- sum of squares
- ANOVA

Plotting means

example with interaction
- nonparallel plotted lines indicate interaction
- significance test determines whether you can conclude the following

three - factor designs
- means that the two-way interactions differ as a function of the level of the third variable
  → usually plot the 2-way interactions separately

V. UNEQUAL SAMPLE SIZES

The Problem of confounding n
- unequal sample sizes can cause confounding

weighted & unweighted means
- to deal with confounding
- means weighted by the factors M_w → ignore effects of other vars = confounding
- unweighted mean = mean of 2 means M_u → control for effect of other vars = eliminate confounding
- SPSS = estimated marginal means
  SAS = least squares means

types of sums of squares
- unequal n → sum of squares total ≠ sum of squares for all sources of var
- Type III sum of squares - when confounded sums of squares are not apportioned to any source of variation
- Type I sum of squares - when confounded sum of squares are all apportioned to sources of variation.
- Type II sum of squares - SS confounded b/n main effects are not apportioned to any source of variation., whereas sums of squares confounded b/n main effects & interactions are apportioned to the main effects

* which Type of SS to use?
- don't use Type 1
- Type II - if theres no interaction → or ANOVA model

VI. Tests supplementing ANOVA

$H_0: \mu_1 = \mu_2 = .. \mu_k$

Main effects
- levels
- significant main
- marginal means

$$Q = \frac{\overbrace{M_i - M_j}^{\text{marginal means}}}{\sqrt{\frac{MSE}{n}}}$$ } equal-sample-size case

MSE → MSE of ANOVA

n → # of scores each mean is based on

* is Tukey HSD test

$$t = \frac{L}{\sqrt{\frac{\sum c_i^2 MSE}{n}}}$$

$$L = \sum \underset{\substack{\uparrow \\ \text{ith coeff}}}{c_i} \underset{\substack{\uparrow \\ \text{marginal mean}}}{M_i}$$

* orthogonal comparisons

INTERACTIONS
- simple effects are different

Describing interactions
1. interaction plot
2. multifactor ANOVA
3. describe difference b/n simple effects

Simple effects tests
- non-significant simple effect does not mean that the simple effect is 0

component of interaction
- orthogonal comparisons
- can test using coefficients

$\boxed{\text{WITHIN SUBJECTS ANOVA}}$
→ comparisons of the same subjects under different conditions
- repeated measures factor
- within subjects design.

- one-Factor Design

- carry over effects
  - symmetric vs asymmetric

- one Between & one within subjects factor

- Assumption of sphericity
  - if violated, leads to increase in type I error rate

Chapter 19: Effect Size

- extent of effect, rather than just presence or absence of the effect

PROPORTIONS:
- compute the difference b/n proportions
- Absolute Risk Reduction (ARR) — the measure of the benefit
  - C - # control
  - T - # treatment

ARR = ~~C~~ C - T

RRR = Relative Risk Reduction:

$$RRR = \frac{C - T}{C \times 100}$$

ODDS Ratio → ratio of odds, not probability

$$N = \frac{1}{ARR}$$

Differenue Between 2 means

- easily interpretable measure of effect size
- wqscale

$$g = \frac{M_1 - M_2}{MSE} \quad \text{Hedges } g$$

$$d = g\sqrt{\frac{N}{N-2}} \quad \text{cohen's } d$$

→ scale free = advantage

interpretational issues

- importance of an effect depends on the context

→ PROPORTION OF VARIANCE EXPLAINED

Anova designs:

measure of effect size: η^2 eta

→ overestimate the variance explained = Biased

$$w^2 = \frac{SSQ_{condition} - (k-1)MSE}{SSQ_{total} + MSE}$$

Factorial designs:

- SS total vs SS Relative

$$w^2 = \frac{SSQ_{effect} - df_{effect} MS_{error}}{SSQ_{total} + MS_{error}}$$

$$w^2_{partial} = \frac{SSQ_{effect} - df_{effect} MS_{error}}{SSQ_{effect} + (N - df_{effect}) MS_{error}}$$

$$R^2_{adjusted} = 1 - \frac{(1-R^2)(N-1)}{N-p-1}$$

Correlational studies

$$\text{prop of var in multiple reg} = \frac{SSQ_{explained}}{SSQ_{total}}$$

→ r^2 in simple reg
R^2 in multiple