Statistics for EMCL week 6

Tamás Biró
Humanities Computing
University of Groningen
t.s.biro@rug.nl

This week

- ANOVA = Analysis of Variance
- Today: one-way ANOVA (M&M 12)
- Advanced stat course: two-way ANOVA (M&M 13)

Inference on means

Sample mean $\bar{x} \longrightarrow$ population mean μ?

- One-sample t-test: is $\mu = \mu_0$?
- Two-sample t-test: is $\mu_1 = \mu_2$?
- More populations, more samples: analysis of variance: is $\mu_1 = \mu_2 = ... = \mu_I$? (ANOVA, "generalized t-test").

Comparing more populations

- Broca's vs. Vernicke's aphasia vs. controls.
- Verbs vs. nouns vs. adjectives vs. function words.
- French vs. English vs. German vs. Dutch.
- Male/N vs. female/N vs. male/V vs. female/V vs. male/A vs. female/A.

Two-way ANOVA

One-way ANOVA: populations $1...I$.

Two-way ANOVA: populations $1...I \times 1...J$.

- Populations can be classified in two ways.
- Mean of responses in two-factor experiments.

Basics of ANOVA

- Populations $1, 2, ...I$.
- Single quantitative variable X on units/cases.
- Interested in pop means $\mu_1, \mu_2, ...\mu_I$ of var X.
- One sample for each population:
  - n_i: size of sample i $(1 \leq i \leq I)$.
  - x_{ij}: case j within sample i $(1 \leq j \leq n_i)$.

H_0 and H_a in ANOVA

- Null hypothesis:
  H_0: $\mu_1 = \mu_2 = ... = \mu_I$.

- Alternative hypothesis:
  not all of the μ_i are equal
  (that is, there exist some i and
  some j such that $\mu_i \neq \mu_j$).

Assumptions of ANOVA

- Fairly Normal distribution per subgroups, no outliers (Normal quantile plot).
- Population standard deviations are equal:
  Instead of performing formal tests: if largest (sample) standard deviation $< 2\times$ smallest (sample) standard deviation.
- Independent observations (watch out for test-retest situations!)

The ANOVA model

- Decompose:
  DATA (total) $=$ FIT (between group) $+$ RESIDUAL (within group).

- F-distribution: reject H_0 if variation among groups large relative to variation within group.

- (F-test for equality of spread/variance M&M 7.3: different from ANOVA, but also employs F-distribution.)

Software output

	Sum of squares	df	Mean square	F	Sig.
Between groups	7.73	3	2.58	11.22	.001
Within groups	3.21	14	.23		
Total	10.94	17			

Reporting ANOVA results

... significant/not significant at $\alpha = 0.05$ level ($F(df1, df2) = ..., p = ...$).

- $df1$: degree of freedom "between groups" (fit, numerator).
- $df2$: degree of freedom "within groups" (residual, denominator).

If ANOVA null hypothesis rejected...

- At least one of the means is different from others. Which one?
- Prior (before data collection) vs. posthoc (after exploratory data analysis).
- (Prior) **contrast**: one-sample t-test with the null hypothesis that $\psi = \sum_{i=1}^{I} a_i \mu_i = 0$ for some a_i's depending on $a\ priori$ hypothesis ($\sum a_i = 0$).

If ANOVA null hypothesis rejected...

- **Multiple comparison**: pairwise comparison of samples i to j.
  - Large I: many comparisons performed.
  - Therefore, reduce α level.
    E.g., Bonferoni: guarantees that the probability of any false rejection no greater than original $\alpha = 5\%$.

ANOVA for exam 1

When to use it:

- Recognize situations in which ANOVA needed.
- What is being tested (null hypothesis, alternative hypothesis).
- Criteria for its use.

ANOVA for exam 2

How to use it:

- Run ANOVA on SPSS.
- Interpret and report results.

Mathematical details (ANOVA model) only for interested.

Next week:

- Non-parametric tests.
- Summary: choice of tests.