Analysis of variance (ANOVA)

Presenter: Steve Baskauf
steve.baskauf@vanderbilt.edu

DISC DIGITAL SCHOLARSHIP AND COMMUNICATIONS

Jean & Alexander Heard LIBRARIES

CodeGraf landing page

- vanderbi.lt/codegraf

One-factor ANOVA

DIGITAL SCHOLARSHIP AND COMMUNICATIONS

Jean & Alexander Heard LIBRARIES

Differences in means for more than 2 groups

- The t-test of means tests whether there is a difference between the means of **two** levels of a factor (i.e. groups of an independent variable).
- Analysis of variance (ANOVA) can test whether there are differences among **three or more** levels (i.e. groups).

How does ANOVA work?

- Variance among observations can be thought of as having two sources:
  - Variance among groups
  - Variance within groups
- The variance within a group is what you would measure for a single level of the variable.
- The variance among groups would be zero if the groups were all the same.
- The statistic F is a ratio that compares the among group variance to the within group variance.

The ANOVA table

```
Response: response
          Df Sum Sq Mean Sq F value    Pr(>F)    
color      2  857.2  428.60  14.813 4.441e-06 ***
Residuals 69 1996.4   28.93                      
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

related to number of levels

derived from variance

P

among

within

related to number of observations

SS/Df

MS_{among}/MS_{within}

- more observations and lower variance within groups makes the denominator of F smaller
- fewer levels and more variance among groups makes the numerator of F larger
- a larger value of F (>1) lowers the value of P

ANOVA syntax in R

- An ANOVA is set up as a linear model:

```
model <- lm(Y ~ X, data = data_frame)
```

- This is like a regression, except that X is discontinuous.

- To generate the ANOVA table for the test:

```
anova(model)
```

- or to output as an ANOVA object:

```
aov(model)
```

Assumptions of ANOVA

1. independence of observations
2. the samples for each level are normally distributed
3. the samples for each level have the same variance (more important)

There are some more technical ones, but these three are basically the same as the t-test of means. Also note:

- ANOVA is sensitive to outliers (so look at a plot)
- ANOVA is robust to violations of #2 if equal sample size and > 10/level
- ANOVA is robust to violations of #3 if equal sample size

(see https://doi.org/10.3758/s13428-017-0918-2)

Comparison of t-test of means to ANOVA

- A single-factor ANOVA with two levels produces exactly the same results as a t-test of means

Post-hoc test for differences among levels

- The Tukey honestly significant difference (HSD) test adjusts the criterion for significance for multiple pairwise comparisons.
- The more comparisons you make, the higher the probability that one of them will be significant by chance (so Tukey HSD is more stringent with more levels present).
- The Tukey-Kramer method applies when sample sizes differ among levels.

```
av <- aov(model)
tukeyHSD(av)
```

Non-parametric alternative to ANOVA

- The Kruskal-Wallis test is a non-parametric alternative to a single-factor ANOVA

```
kruskal.test(Y ~ X, data = data_frame)
```

Two-factor ANOVA

DIGITAL SCHOLARSHIP AND COMMUNICATIONS

Jean & Alexander Heard LIBRARIES

Two-factor ANOVA minefield

- There are many variants including fixed effects, random effects, mixed model, nested, repeated measures, and Types I, II, and III.
- Different software packages calculate the statistics differently, resulting in different results for the "same" test.
- Some statisticians doubt the validity of calculating P in some circumstances.
- Some statisticians believe that there are better alternatives to two-factor ANOVA.
- However, if the experimental design is simple and balanced, there are less likely to be problems.

"Full factorial" experiment

- A **full factorial** experiment has every combination of every level of every factor.
- The factors are two **independent variables** (X's), a.k.a. "effects"
- Both effects are **fixed** since we control their presence or absence
- Example
  - factors: soap and antimicrobial agent (triclosan)
  - levels of each factor: present, absent

FACTORS		soap	
	LEVELS	present	absent
antimicrobial	present	has soap and antimicrobial	has antimicrobial but no soap
	absent	has soap but no antimicrobial	neither soap nor antimicrobial

"Tidy" data setup

- Two columns of factors (soap, triclosan)
- Each factor has levels: "yes" and "no"
- The counts are the **dependent variable** (Y)
- This is a **balanced design** (equal sample size for each combination of levels)

(note: these are fake data)

soap	triclosan	counts
yes	yes	3200
yes	yes	4300
yes	yes	1600
yes	yes	3800
yes	yes	2500
yes	no	3900
yes	no	1200
yes	no	2200
yes	no	3400
yes	no	2300
no	yes	1600
no	yes	2400
no	yes	1900
no	yes	1300
no	yes	2100
no	no	1900
no	no	1800
no	no	1100
no	no	2600
no	no	2000

Questions we want to investigate

1. Does soap have an effect?
2. Does triclosan have an effect?
3. Is there an interaction between soap and triclosan?

Notes:

- "soap" and "triclosan" are called the **main effects** (vs. the interaction).
- If there is an interaction, then asking the first two questions doesn't make sense.

Form of the model

Note: there are several ways to set up ANOVAs in R. This is just one.

```
aov(Y ~ X1 + X2 + X1:X2)
```

Example:

```
aov(counts ~ soap + triclosan + soap:triclosan, data = dataframe)
```

Shortcut for all factors and interaction:

```
aov(counts ~ soap * triclosan, data = dataframe)
```

ANOVA table for two factors

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
soap	1	4704500	4704500	6.863	0.0186	*
triclosan	1	264500	264500	0.386	0.5432	
soap:triclosan	1	312500	312500	0.456	0.5092	
Residuals	16	10968000	685500			

- The method of calculating these statistics is complex, although the interpretation is similar to single factor.
- In this case, the **`soap:triclosan`** interaction is not significant, so it makes sense to examine the main effects.

Blocking (random effects)

DISC DIGITAL SCHOLARSHIP AND COMMUNICATIONS

Jean & Alexander Heard LIBRARIES

Random effects

- A **random** (vs. fixed) effect is not controlled by the experimenter.
  - Example: the cockroaches and their situation vary in random ways.
  - Other examples: location in an experimental site, days on which an experiment is conducted.
- A **block** is an experimental unit onto which all of the fixed effects are applied.
  - Example: each cockroach eye had every color of light applied to it
- Including random effects:
  - allows us to assess their influence
  - removes their variability from the residuals (increases power)
- Statistics for random effects in ANOVA are **calculated differently** from fixed effects.

Experimental design

- The levels of the random factor don't have any significance experimentally, the are simply **grouping variables**.
- Each cell has only one value (**no replication**)
- Because there is no replication, the **interaction can't be determined**.

FACTOR		color (fixed)		
	LEVELS	red	green	blue
block (random)	a	one voltage	one voltage	one voltage
	b	one voltage	one voltage	one voltage
	c	one voltage	one voltage	one voltage
	e	one voltage	one voltage	one voltage
	f	one voltage	one voltage	one voltage
	…			

Form of the model

`lme4` package is for constructing linear mixed models: **`lmer()`** vs. **`lm()`**

```
lmer(Y ~ X1 + (1 | X2))
```

Example:

```
mixed_model <- lmer(response ~ color + (1 | block), data = erg_dataframe)
```

- The random effect is specified by **`(1 | X2)`**
- There is no interaction effect (no replication).
- The **`lme4`** package does not calculate *P*. **`lmerTest`** is a wrapper that adds *P*

Comparison of paired t-test to ANOVA

- A two-factor ANOVA with blocking, two levels, and a balanced design produces the same result as a paired t-test.