Unit 4: Inference for numerical data
4. ANOVA

Sta 101 - Fall 2015

Duke University, Department of Statistical Science

Dr. Çetinkaya-Rundel

Slides posted at *http://bit.ly/sta101_f15*

- PS 4 due Friday
- PA 4 due Sunday
- RA 5 next Monday

NEWS FLASH!
Jelly beans rumored to cause acne!!!

How would you check this rumor? Imagine that doctors can assign an "acne score" to patients on a 0-100 scale.

- What would your research question be?
- How would you conduct your study?
- What statistical test would you use?

http://imgs.xkcd.com/comics/significant.png

Clicker question

Suppose $\alpha = 0.05$. What is the probability of making a Type 1 error and rejecting a null hypothesis like

$$H_0 : \mu_{\text{purple jelly bean}} - \mu_{\text{placebo}} = 0$$

when it is actually true?

(a) 1%
(b) 5%
(c) 36%
(d) 64%
(e) 95%

Clicker question

Suppose we want to test 20 different colors of jelly beans versus a placebo with hypotheses like

$$H_0 : \mu_{\text{purple jelly bean}} - \mu_{\text{placebo}} = 0$$
$$H_0 : \mu_{\text{brown jelly bean}} - \mu_{\text{placebo}} = 0$$
$$H_0 : \mu_{\text{peach jelly bean}} - \mu_{\text{placebo}} = 0$$
$$\ldots$$

and we use $\alpha = 0.05$ for each of these tests. What is the probability of making at least one Type 1 error in these 20 independent tests?

(a) 1%
(b) 5%
(c) 36%
(d) 64%
(e) 95%

ANOVA tests for some difference in means of many different groups

Null hypothesis:

$$H_0 : \mu_{\text{placebo}} = \mu_{\text{purple}} = \mu_{\text{brown}} = \ldots = \mu_{\text{peach}} = \mu_{\text{orange}}.$$

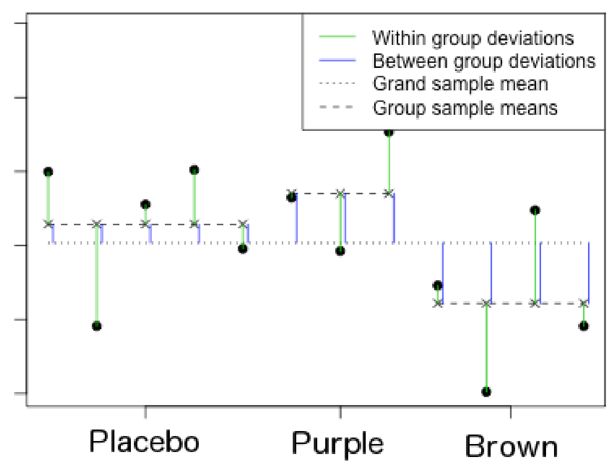
Clicker question

Which of the following is a correct statement of the alternative hypothesis?

(a) For any two groups, including the placebo group, no two group means are the same.
(b) For any two groups, not including the placebo group, no two group means are the same.
(c) Amongst the jelly bean groups, there are at least two groups that have different group means from each other.
(d) Amongst all groups, there are at least two groups that have different group means from each other.

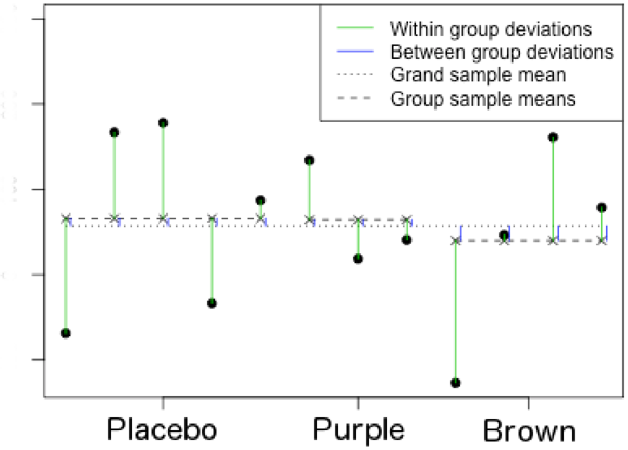
ANOVA compares between group variation to within group variation

$$\sum |^2 \Big/ \sum |^2$$



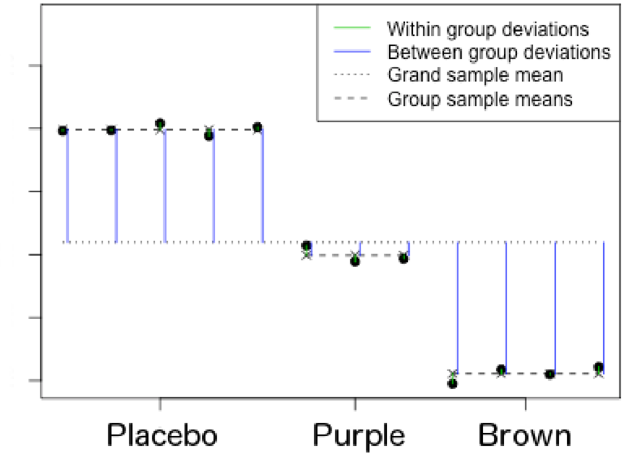
Relatively large WITHIN group variation: little apparent difference

$$\sum |^2 \Big/ \sum |^2$$



Relatively large BETWEEN group variation: there may be a difference

$$\sum |^2 \Big/ \sum |^2$$



For historical reasons, we use a modification of this ratio called the F-statistic:

$$F = \frac{\sum |^2 \ / \ (k-1)}{\sum |^2 \ / \ (n-k)} = \frac{MSG}{MSE}$$

k: # of groups; n: # of obs.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Between groups	k-1	$\sum \|^2$	MSG	F_{obs}	p_{obs}
Within groups	n-k	$\sum \|^2$	MSE		
Total	n-1	$\sum(\| + \|)^2$			

To identify which means are different, use t-tests and the Bonferroni correction

- If the ANOVA yields a significant results, next natural question is: "Which means are different?"
- Use t-tests comparing each pair of means to each other,
  - with a common variance (*MSE* from the ANOVA table) instead of each group's variances in the calculation of the standard error,
  - and with a common degrees of freedom (df_E from the ANOVA table)
- Compare resulting p-values to a modified significance level

$$\alpha^\star = \frac{\alpha}{K}$$

where K is the total number of pairwise tests

Application exercise: 4.4 ANOVA

See the course webpage for details.

1. Comparing many means requires care
2. ANOVA tests for <u>some</u> difference in means of many different groups
3. ANOVA compares between group variation to within group variation
4. To identify which means are different, use t-tests and the Bonferroni correction