Unit 4: Inference for numerical data
3. ANOVA

Sta 101 - Spring 2015

Duke University, Department of Statistical Science

March 2, 2015

Dr. Windle

Slides posted at http://bitly.com/windle2

NEWS FLASH!

Jelly beans rumored to affect 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

1. It is difficult to simultaneously compare many groups.

Suppose $\alpha = 0.05$.

What is the probability of correctly failing to reject

$$H_0 : \mu_{\text{purple}} = \mu_{\text{placebo}} \ ?$$

Clicker question

If no color of Jelly bean has any link to acne, what is the probability of making at least one type I error in the 20 trials?

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

2. ANOVA is useful for testing if there is some difference between the means of many different groups.

Null hypothesis for "F-Test" (the test associated with ANOVA):

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

Clicker question

Which of the following is a correct version 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.
(d) Amongst all groups, there are at least two groups that have different group means.

2. ANOVA is useful for testing if there is some difference between the means of many different groups.

The practical implication of this alternative is: ``At least one color of jelly bean is linked to acne.''

At least two of the group means are not the same:

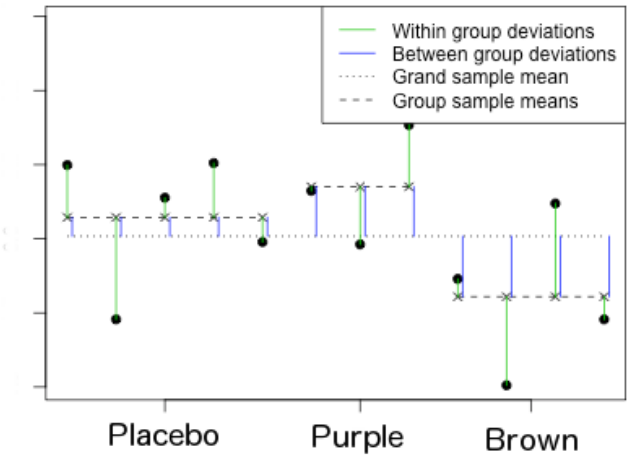
1. $\mu_{\text{placebo}} \neq \mu_{\text{color}}$ for some color of jelly bean, or
2. $\mu_A \neq \mu_B$ for two colors, A and B.

   Then

   1. $\mu_A \neq \mu_{\text{placebo}}$, or
   2. $\mu_A = \mu_{\text{placebo}}$. Thus, $\mu_B \neq \mu_A = \mu_{\text{placebo}}$.

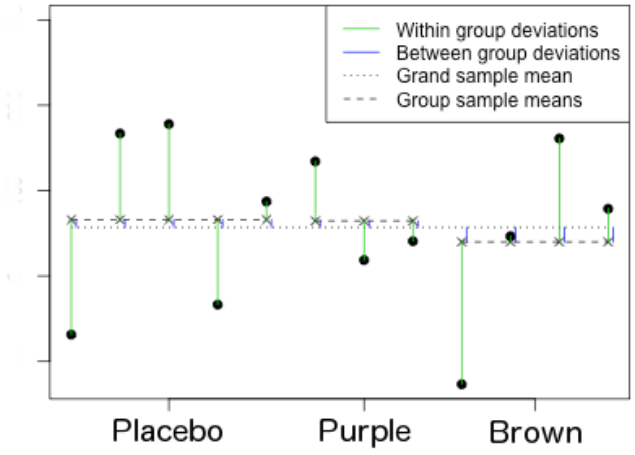
3. The test is based on comparing between group to within group variation.

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



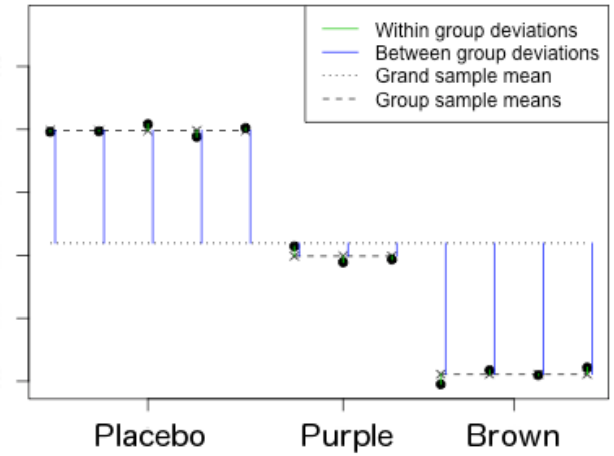
Relatively large WITHIN group variation: little apparent difference

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



Relatively large BETWEEN group variation: there may be a difference

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



3. The F-test is based on comparing between group to within group variation

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

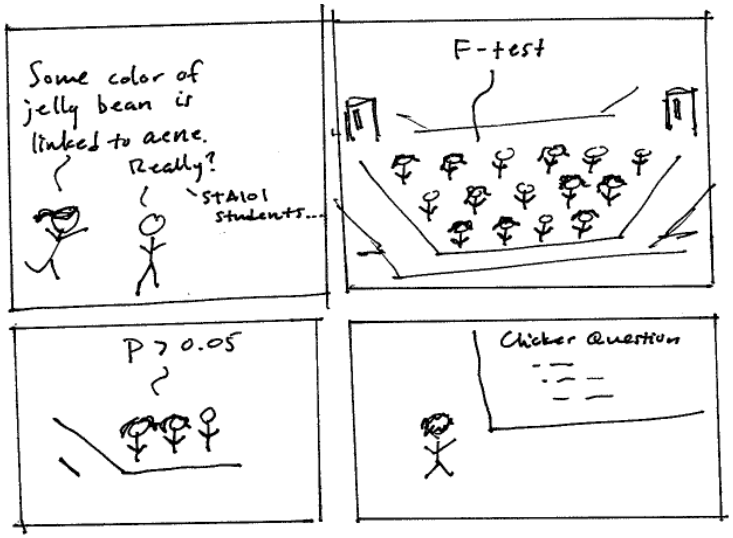
$$F = \frac{\sum |^2 \ / \ (j-1)}{\sum |^2 \ / \ (n-j)} \quad = \quad \frac{MSG}{MSE}.$$

j: # of groups; *n*: # of obs.

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

$$p_{obs} = p(W > F_{obs} \mid H_0) = p(W > F_{obs} \mid W \sim \text{F-dist}_{j-1,n-j})$$

`bitly.com/dist_calc`



Say $\alpha = 0.05$.

Clicker question

What is the most accurate statement of the results?

(a) At least one color of jelly bean is linked to acne.
(b) At least one color of jelly bean is not linked to acne.
(c) There is little evidence that any color of jelly bean is linked to acne.
(d) Jelly beans definitely do not cause acne.

Clicker question

For the *F*-test, what is the probability of incorrectly rejecting the null?

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

Application exercise: 4.4 ANOVA - Pt 1

See the course webpage for details.

Summary of main ideas

1. It is difficult to simultaneously compare many groups.
2. ANOVA is useful for testing if there is some difference between the means of many different groups.
3. The test is based on comparing between group to within group variation.