Unit 4: Inference for numerical data

3. ANOVA

Sta 104 - Summer 2015

Duke University, Department of Statistical Science

June 3, 2015

Dr. Çetinkaya-Rundel

Slides posted at http://bit.ly/sta104su15

Outline

1. Housekeeping

2. 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

3. Summary

Announcements

1. Housekeeping

2. 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

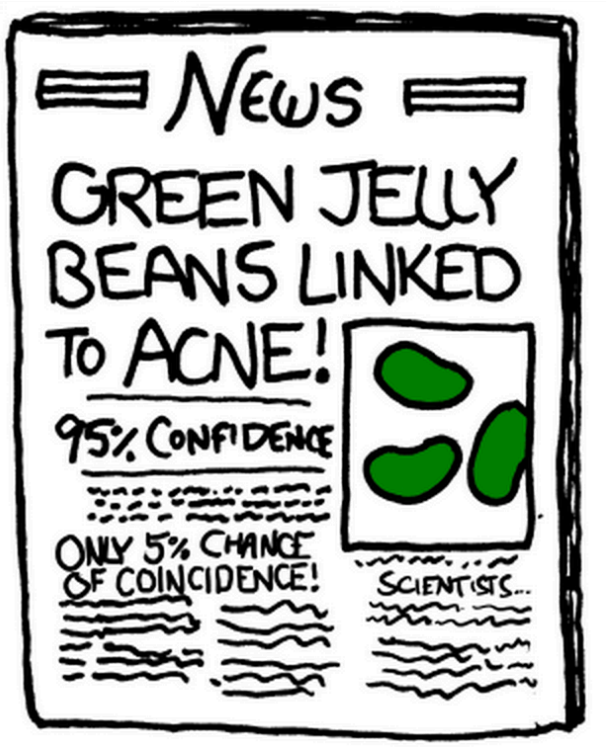
3. Summary

1. Housekeeping

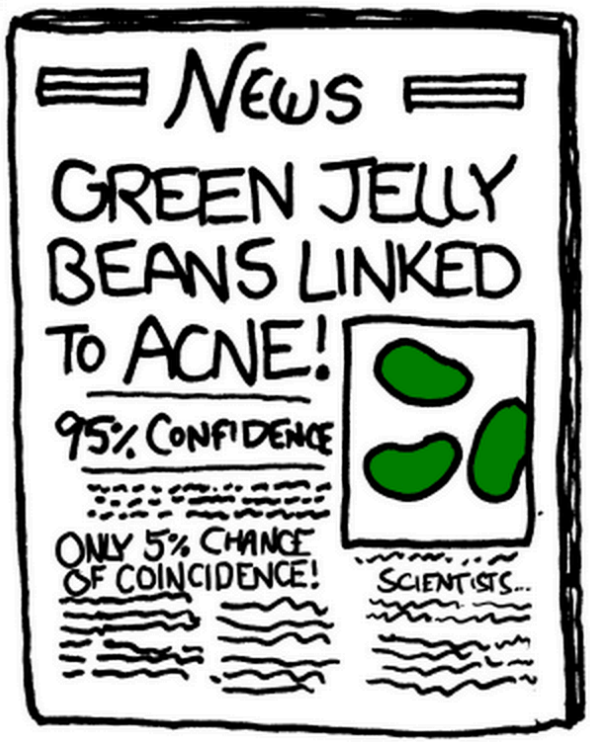
2. Main ideas
    1. It is difficult to simultaneously compare many groups
    2. ANOVA is useful for testing if there is some difference between the means of many different groups
    3. The test is based on comparing between group to within group variation

3. Summary

1. It is difficult to simultaneously compare many groups

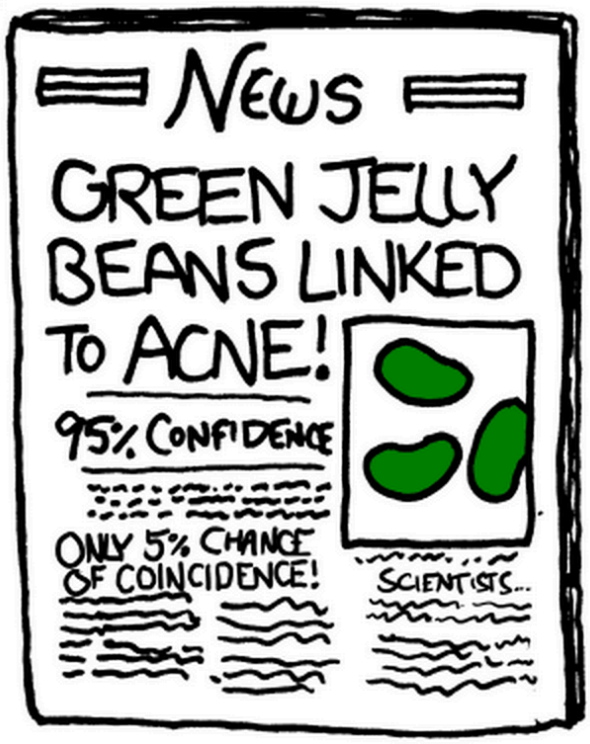


1. It is difficult to simultaneously compare many groups



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

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



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

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

Use an independent samples t-test:

$$H_0 : \mu_{\textit{green jelly beans}} = \mu_{\textit{placebo}}$$

$$H_A : \mu_{\textit{green jelly beans}} \neq \mu_{\textit{placebo}}$$

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

Suppose $\alpha = 0.05$.

What is the probability of rejecting the following null hypothesis when in fact it is true?

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

Suppose $\alpha = 0.05$.

What is the probability of rejecting the following null hypothesis when in fact it is true?

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

P(Type 1 Error) = 0.05

Suppose $\alpha = 0.05$.

Clicker question

If all the tests are independent and 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%

Suppose $\alpha = 0.05$.

Clicker question

If all the tests are independent and 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%* $\rightarrow 1 - 0.95^{20} \approx 0.64$

(d) 95%

Outline

1. Housekeeping

2. 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

3. Summary

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.

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.*

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

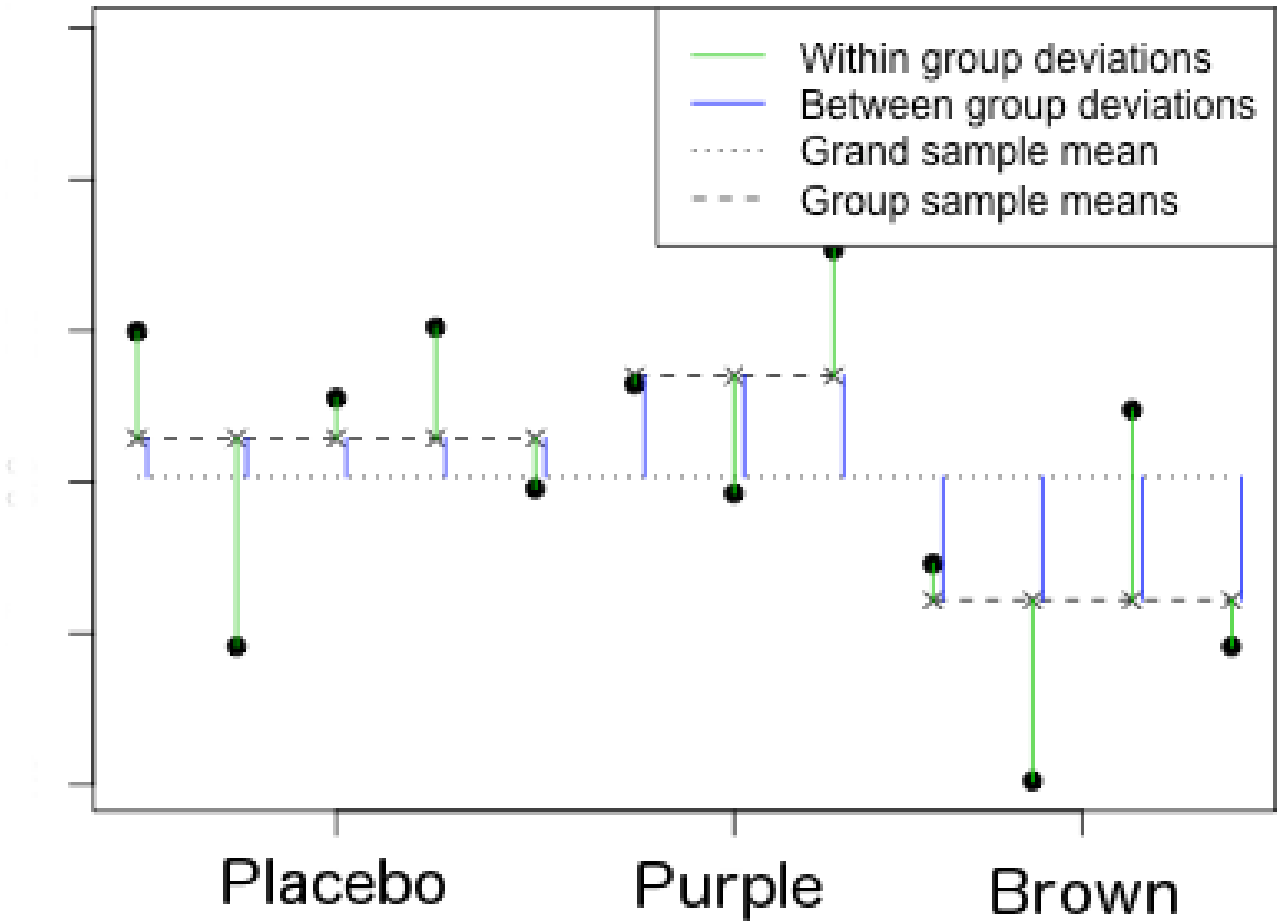
1. Housekeeping

2. 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

3. Summary

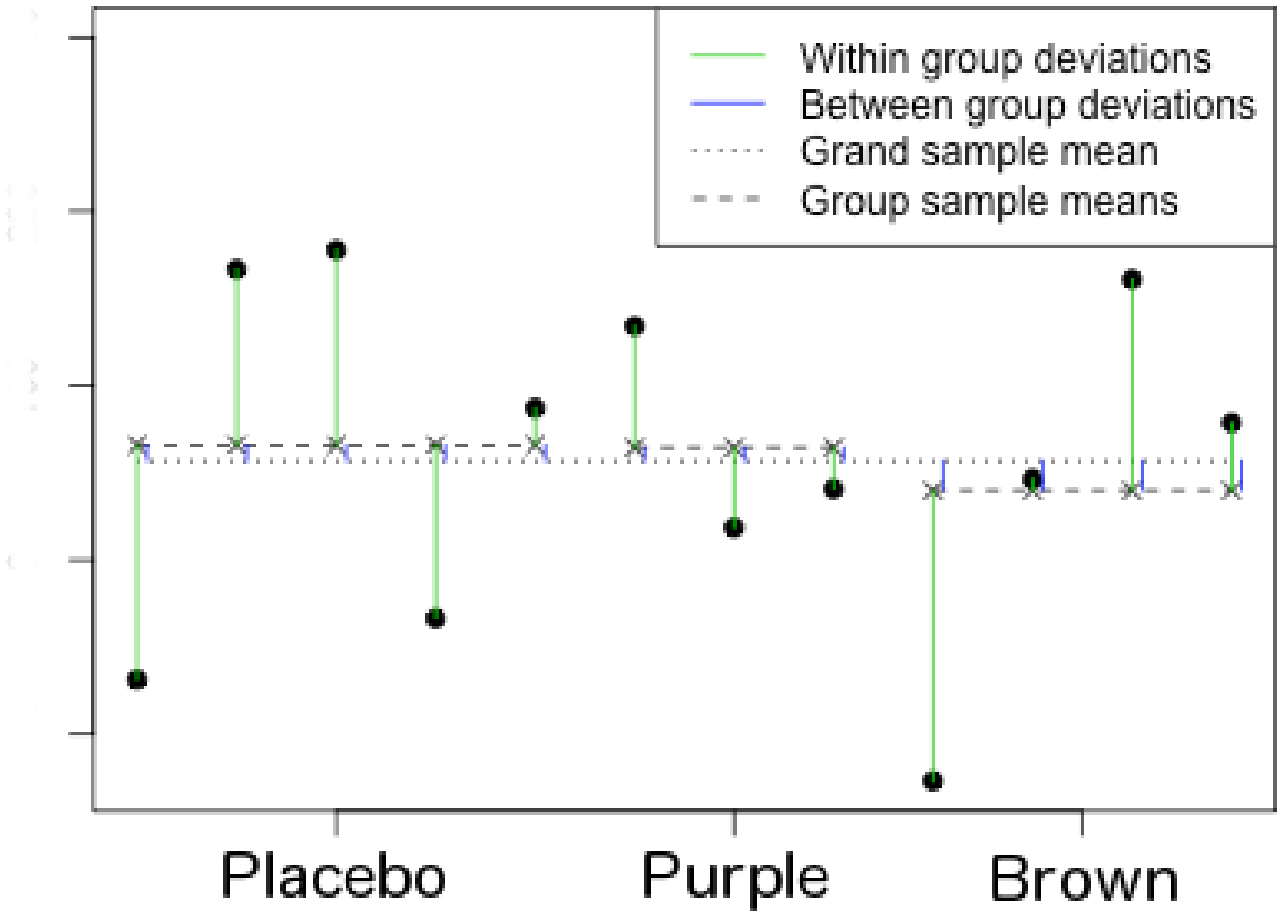
3. The test compares between group variation to within group variation

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



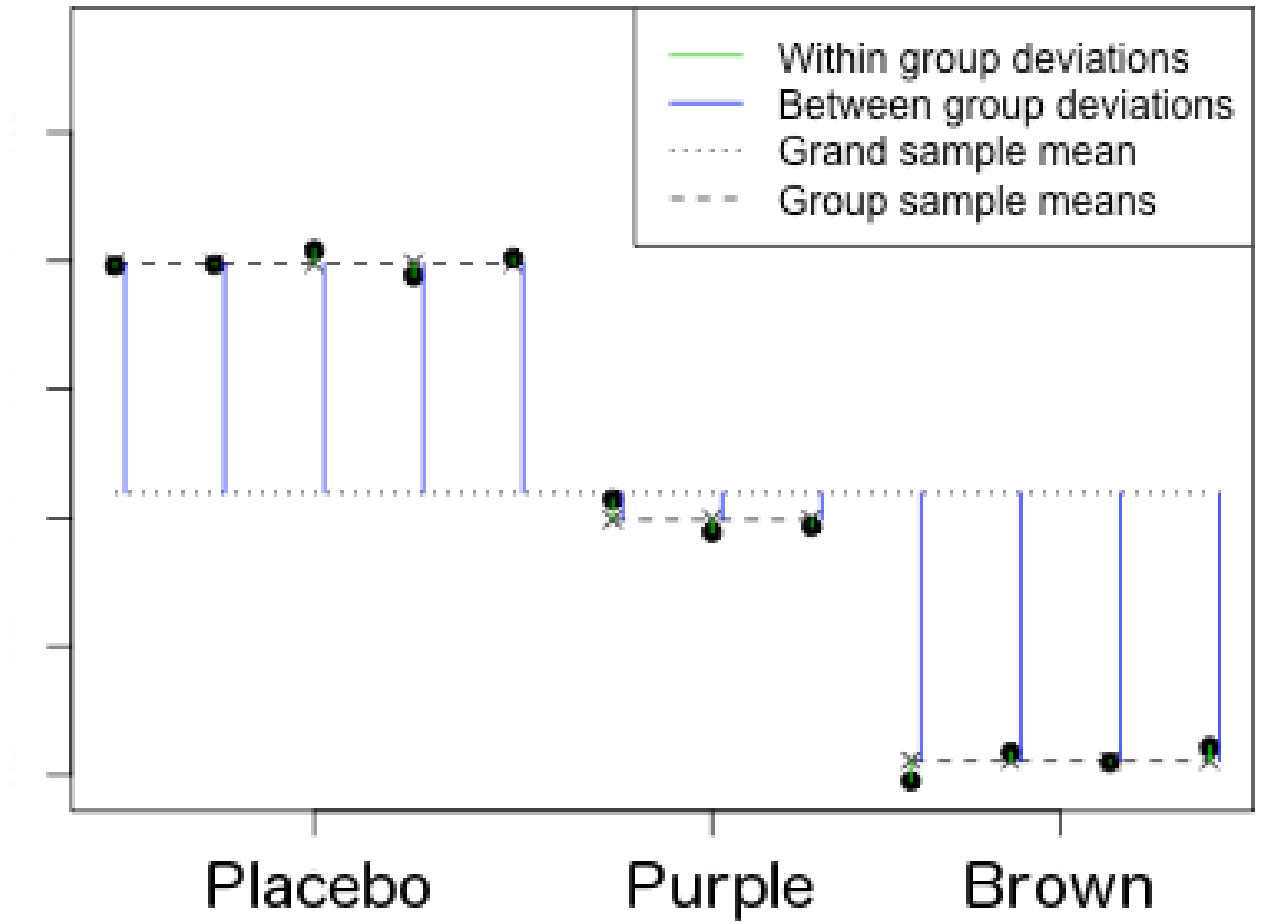
Relatively large WITHIN group variation: little apparent difference

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



Relatively large BETWEEN group variation: there may be a difference

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



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

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

where k is the # of groups and n is the # of observations

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

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

where k is the # of groups and n is the # of observations

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

Clicker question

The p-value for the F-test is 0.045, and $\alpha = 0.05$. 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

The p-value for the F-test is 0.045, and $\alpha = 0.05$. 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 with $\alpha = 0.05$, what is the probability of incorrectly rejecting the null?

(a) 5%

(b) 36%

(c) 64%

(d) 95%

Clicker question

For the F-test with $\alpha = 0.05$, what is the probability of incorrectly rejecting the null?

(a) *5%*

(b) 36%

(c) 64%

(d) 95%

Application exercise: 4.5 ANOVA - Pt 1

See the course webpage for details.

Outline

1. Housekeeping

2. 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

3. Summary

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