Unit 4: Inference for numerical data

3. ANOVA

Sta 101 - Spring 2015

Duke University, Department of Statistical Science

March 4, 2015

Dr. Windle

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

Outline

1. Housekeeping

2. Main Ideas
    1. If you want to test many hypotheses simulteneously, use the Bonferroni correction.
    2. You can use another version of the F-test to compare grouping by 2 variables vs. grouping by 1 variable

3. Summary

Announcements

Outline

1. Housekeeping

2. Main Ideas

    1. If you want to test many hypotheses simulteneously, use the Bonferroni correction.

    2. You can use another version of the F-test to compare grouping by 2 variables vs. grouping by 1 variable

3. Summary

Outline

1. Housekeeping

2. Main Ideas

    1. If you want to test many hypotheses simulteneously, use the Bonferroni correction.

    2. You can use another version of the F-test to compare grouping by 2 variables vs. grouping by 1 variable

3. Summary

An example by James G. Scott:

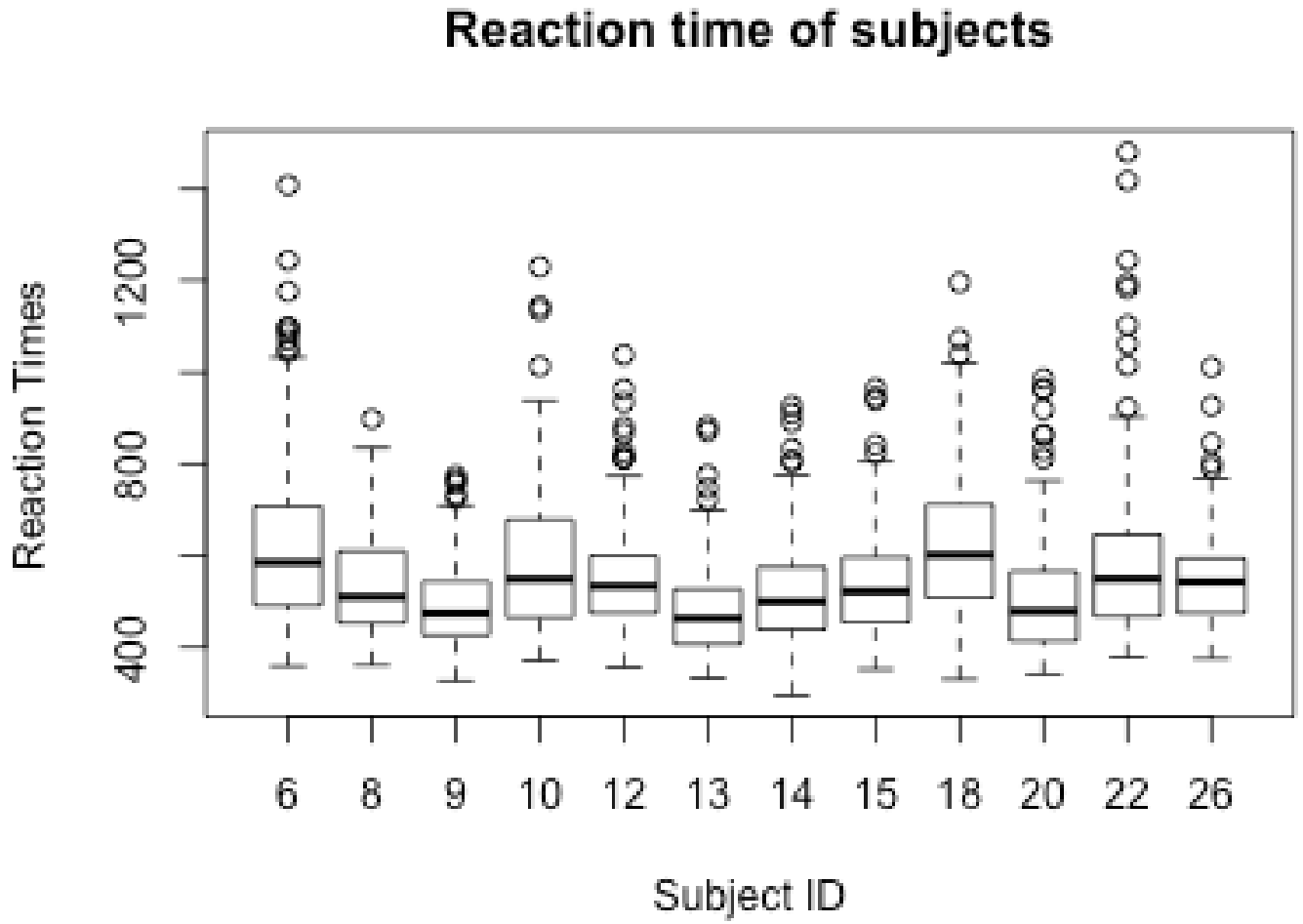
> *[These data are] from an experiment run by a British video-game manufacturer in an attempt to calibrate the level of difficulty of certain tasks in the video game. Subjects in this experiment were presented with a simple "Where's Waldo?"-style visual scene. The subjects had to find a number (1 or 2) floating somewhere in the scene, to identify the number, and to press the corresponding button as quickly as possible. The response variable is their reaction time.*

Obs	Subject	PictureTarget.RT	Littered	FarAway
1	10	635	0	0
2	10	1144	0	0
3	10	570	0	0
4	10	589	0	0
5	10	754	0	0
6	10	601	0	0

From JGS:

- `PictureTarget.RT`: the subject's reaction time in milliseconds.
- `Subject`: a numerical identifier for the subject undergoing the test.
- `FarAway`: was the number to be identified far away (1) or near (0) in the visual scene?
- `Littered`: the British way of saying whether the scene was cluttered (1) or mostly free of clutter (0).

Do some subjects in the study have different mean reaction times?



Number of observations $n = 1920$.

ANOVA Table

	Df	Sum Sq	Mean Sq	Fvalue	Pr(>F)
Subject	??	4060822	369166	20.5	2.2e-16
Residuals	??	35129401	18412		

Clicker question

What are the degrees of freedom?

(a) 1 and 1909
(b) 11 and 1908
(c) 11 and 1909
(d) 12 and 1908

Number of observations $n = 1920$.

ANOVA Table

	Df	Sum Sq	Mean Sq	Fvalue	Pr(>F)
Subject	??	4060822	369166	20.5	2.2e-16
Residuals	??	35129401	18412		

Clicker question

What are the degrees of freedom?

(a) 1 and 1909
(b) *11 and 1908*
(c) 11 and 1909
(d) 12 and 1908

(Assume $\alpha = 0.05$.)

Clicker question

What is the most appropriate conclusion?

(a) There is no evidence that the subjects have different mean reaction times.

(b) There is no evidence that some of the subjects have the same mean reaction times.

(c) Some pairs of subjects have different mean reaction times.

(d) All paris of subjects have different mean reaction times.

(Assume $\alpha = 0.05$.)

Clicker question

What is the most appropriate conclusion?

(a) There is no evidence that the subjects have different mean reaction times.

(b) There is no evidence that some of the subjects have the same mean reaction times.

(c) *Some pairs of subjects have different mean reaction times.*

(d) All paris of subjects have different mean reaction times.

How do we determine which subjects have a different mean reaction time than Subject 6?

How do we determine which subjects have a different mean reaction time than Subject 6?

The Bonferroni correction.

Goal: test 11 different hypotheses:

$$
\begin{aligned}
01. \quad & H_0 : \mu_{\text{S06}} = \mu_{\text{S08}} \\
& H_A : \mu_{\text{S06}} \neq \mu_{\text{S08}} \\
02. \quad & H_0 : \mu_{\text{S06}} = \mu_{\text{S09}} \\
& H_A : \mu_{\text{S06}} \neq \mu_{\text{S09}} \\
& \ldots \\
11. \quad & H_0 : \mu_{\text{placebo}} = \mu_{\text{S26}} \\
& H_A : \mu_{\text{placebo}} \neq \mu_{\text{S26}}
\end{aligned}
$$

AND keep the Type I error rate at or below the significance level.

1. If you want to test many hypotheses simulteneously, use the Bonferroni correction.

Bonferroni correction:

- Target type I error rate: α.
- Number of null/alt hypotheses to be tested using the same data set: K
- If you set the significance level for each test to be

$$\alpha^* = \alpha / K,$$

then the probability of making one or more type I errors is $\leq \alpha$.

1. If you want to test many hypotheses simulteneously, use the Bonferroni correction.

Hyp	p-value
01	4.27e-09
02	< 2e-16
03	0.00368
04	5.68e-07
05	< 2e-16
06	1.82e-13
07	1.11e-09
08	0.61587
09	1.42e-14
10	0.02332
11	1.91e-07

Clicker question

Which null hypotheses should we reject at $\alpha = 0.05$?

(a) 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11

(b) 1, 2, 3, 4, 5, 6, 7, 9, 11

(c) 1, 2, 4, 5, 6, 7, 9, 11

(d) 1, 2, 5, 6, 7, 9, 11

1. If you want to test many hypotheses simulteneously, use the Bonferroni correction.

Hyp	p-value
01	4.27e-09
02	< 2e-16
03	0.00368
04	5.68e-07
05	< 2e-16
06	1.82e-13
07	1.11e-09
08	0.61587
09	1.42e-14
10	0.02332
11	1.91e-07

Clicker question

Which null hypotheses should we reject at $\alpha = 0.05$?

(a) 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11

(b) *1, 2, 3, 4, 5, 6, 7, 9, 11*

(c) 1, 2, 4, 5, 6, 7, 9, 11

(d) 1, 2, 5, 6, 7, 9, 11

1. If you want to test many hypotheses simulteneously, use the Bonferroni correction.

Application exercise: 4.5 ANOVA - Pt 2

See the course webpage for details.

Outline

1. Housekeeping

2. Main Ideas

    1. If you want to test many hypotheses simulteneously, use the Bonferroni correction.

    2. You can use another version of the F-test to compare grouping by 2 variables vs. grouping by 1 variable

3. Summary

You can use another version of the F-test to compare grouping by 2 variables vs. grouping by 1 variable

A new research question:

Does "litter" in the image change some subjects' reaction times?

$$
\begin{aligned}
01.\quad & H_0 : \mu_{\text{S06 \& Litter}} = \mu_{\text{S06 \& No Litter}} \\
& H_A : \mu_{\text{S06 \& Litter}} \neq \mu_{\text{S06 \& No Litter}} \\
02.\quad & H_0 : \mu_{\text{S08 \& Litter}} = \mu_{\text{S08 \& No Litter}} \\
& H_A : \mu_{\text{S08 \& Litter}} \neq \mu_{\text{S08 \& No Litter}} \\
& \ldots \\
12.\quad & H_0 : \mu_{\text{S26 \& Litter}} = \mu_{\text{S26 \& No Litter}} \\
& H_A : \mu_{\text{S26 \& Litter}} \neq \mu_{\text{S26 \& No Litter}}
\end{aligned}
$$

Same problem as before... multiple hypotheses.

You can use another version of the F-test to compare grouping by 2 variables vs. grouping by 1 variable

Null hypothesis: litter does not change the mean reaction time of anyone.

$$
\begin{aligned}
H_0 : \mu_{\text{S06 \& Litter}} &= \mu_{\text{S06 \& No Litter}} \\
\text{and } \mu_{\text{S08 \& Litter}} &= \mu_{\text{S08 \& No Litter}} \\
\text{and } &\ldots \\
\text{and } \mu_{\text{S26 \& Litter}} &= \mu_{\text{S26 \& No Litter}}
\end{aligned}
$$

You can use another version of the F-test to compare grouping by 2 variables vs. grouping by 1 variable

If litter does not change the mean reaction time of anyone, then, for instance,

$$\mu_{\text{S06 \& Litter}} = \mu_{\text{S06 \& No Litter}} = \mu_{\text{S06}}. \tag{1}$$

You can use another version of the F-test to compare grouping by 2 variables vs. grouping by 1 variable

If litter does not change the mean reaction time of anyone, then, for instance,

$$\mu_{\text{S06 \& Litter}} = \mu_{\text{S06 \& No Litter}} = \mu_{\text{S06}}. \tag{1}$$

Think about within group SS:

$$SSE = \sum_{i=1}^{n} (y_i - \mu_{\text{group that } y_i \text{ is in}})^2.$$

You can use another version of the F-test to compare grouping by 2 variables vs. grouping by 1 variable

If litter does not change the mean reaction time of anyone, then, for instance,

$$\mu_{\text{S06 \& Litter}} = \mu_{\text{S06 \& No Litter}} = \mu_{\text{S06}}. \tag{1}$$

Think about within group SS:

$$SSE = \sum_{i=1}^{n} (y_i - \mu_{\text{group that } y_i \text{ is in}})^2.$$

If (1) is true for every subject, then

$$SSE_{\text{subject}} = SSE_{\text{subject \& litter}}.$$

You can use another version of the F-test to compare grouping by 2 variables vs. grouping by 1 variable

To see if there is a difference between some groups, look at percent increase in within group variation when excluding litter as an explanatory variable:

$$\frac{SSE_{\text{subject}} - SSE_{\text{subject \& litter}}}{SSE_{\text{subject \& litter}}}.$$

- Small: no difference for any subject.
- Large: a difference for some subject.

You can use another version of the F-test to compare grouping by 2 variables vs. grouping by 1 variable

To see if there is a difference between some groups, look at percent increase in within group variation when excluding litter as an explanatory variable:

$$F = \frac{(SSE_{\text{subject}} - SSE_{\text{subject \& litter}})/(j_{\text{subject \& litter}} - j_{\text{subject}})}{SSE_{\text{subject \& litter}}/(n - j_{\text{subject \& litter}})}.$$

- Small: no difference for any subject.
- Large: a difference for some subject.

You can use another version of the F-test to compare grouping by 2 variables vs. grouping by 1 variable

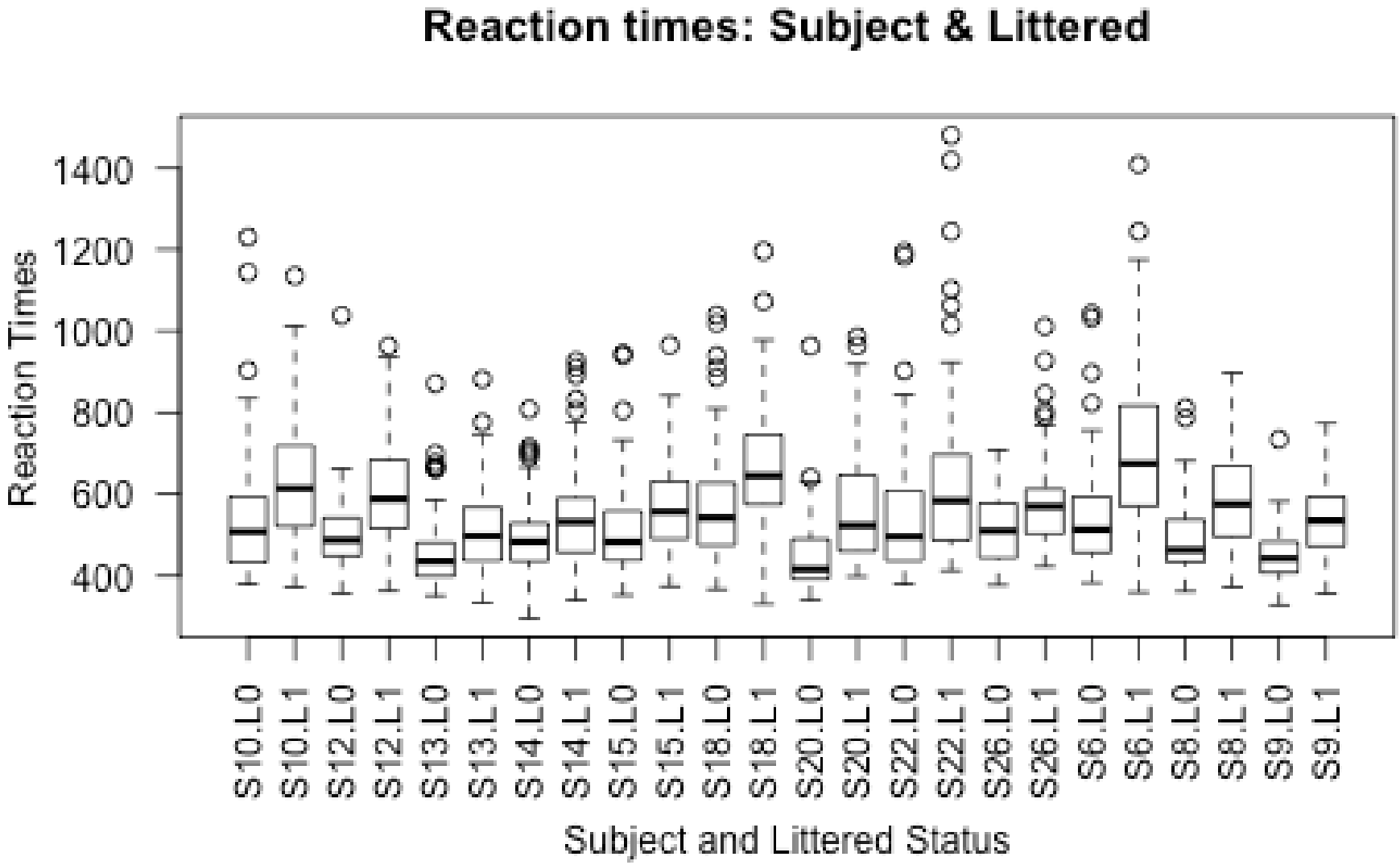
To see if there is a difference between some groups, look at percent increase in within group variation when excluding litter as an explanatory variable:

$$F = \frac{(SSE_{\text{subject}} - SSE_{\text{subject \& litter}})/(j_{\text{subject \& litter}} - j_{\text{subject}})}{SSE_{\text{subject \& litter}}/(n - j_{\text{subject \& litter}})}.$$

- Small: no difference for any subject.
- Large: a difference for some subject.

n: number of obs.
j_{subject}: number of groups by subject.
$j_{\text{subject \& litter}}$: number of groups by subject and litter.

You can use another version of the F-test to compare grouping by 2 variables vs. grouping by 1 variable



You can use another version of the F-test to compare grouping by 2 variables vs. grouping by 1 variable

$$df_1 = j_{\text{subject \& litter}} - j_{\text{subject}}$$
$$df_2 = n - j_{\text{subject \& litter}}.$$

$$F_{obs} = \frac{(SSE_{\text{subject}} - SSE_{\text{subject \& litter}})/(j_{\text{subject \& litter}} - j_{\text{subject}})}{SSE_{\text{subject \& litter}}/(n - j_{\text{subject \& litter}})}.$$

$$p_{obs} = P(W > F_{obs} \mid H_0) = P(W > F_{obs} \mid W \sim \text{F-dist}_{df_1, df_2}).$$

Clicker question

$SSE_{\text{subject \& litter}} = 30935621$; $SSE_{\text{subject}} = 35129401$. There are 1920 obs. Do you reject at the $\alpha = 10^{-5}$ level?

(A) Yes (b) No

You can use another version of the F-test to compare grouping by 2 variables vs. grouping by 1 variable

$$df_1 = j_{\text{subject \& litter}} - j_{\text{subject}}$$
$$df_2 = n - j_{\text{subject \& litter}}.$$

$$F_{obs} = \frac{(SSE_{\text{subject}} - SSE_{\text{subject \& litter}})/(j_{\text{subject \& litter}} - j_{\text{subject}})}{SSE_{\text{subject \& litter}}/(n - j_{\text{subject \& litter}})}.$$

$$p_{obs} = P(W > F_{obs} \mid H_0) = P(W > F_{obs} \mid W \sim \text{F-dist}_{df_1, df_2}).$$

Clicker question

$SSE_{\text{subject \& litter}} = 30935621$; $SSE_{\text{subject}} = 35129401$. There are 1920 obs. Do you reject at the $\alpha = 10^{-5}$ level?

(A) *Yes* (b) No

You can use another version of the F-test to compare grouping by 2 variables vs. grouping by 1 variable

Clicker question

What is the best interpretation of the result?

There is evidence that

(a) some of the subject's reaction times change when the image is littered.

(b) many subjects reaction times change when the image is littered, since the F-value is large.

(c) all of the subject's reaction times change when the image is littered.

What would you do to test which subjects' reaction times change when the image is littered?

You can use another version of the F-test to compare grouping by 2 variables vs. grouping by 1 variable

Clicker question

What is the best interpretation of the result?

There is evidence that

(a) *some of the subject's reaction times change when the image is littered.*

(b) many subjects reaction times change when the image is littered, since the F-value is large.

(c) all of the subject's reaction times change when the image is littered.

What would you do to test which subjects' reaction times change when the image is littered?

Outline

1. Housekeeping

2. Main Ideas
    1. If you want to test many hypotheses simulteneously, use the Bonferroni correction.
    2. You can use another version of the F-test to compare grouping by 2 variables vs. grouping by 1 variable

3. Summary

1. If you want to test many hypotheses simulteneously, use the Bonferroni correction.
2. You can use another version of the F-test to compare grouping by 2 variables vs. grouping by 1 variable