# Unit 4: Inference for Numerical Data
# Lecture 3: Comparing many means via ANOVA

Statistics 101

Mine Çetinkaya-Rundel

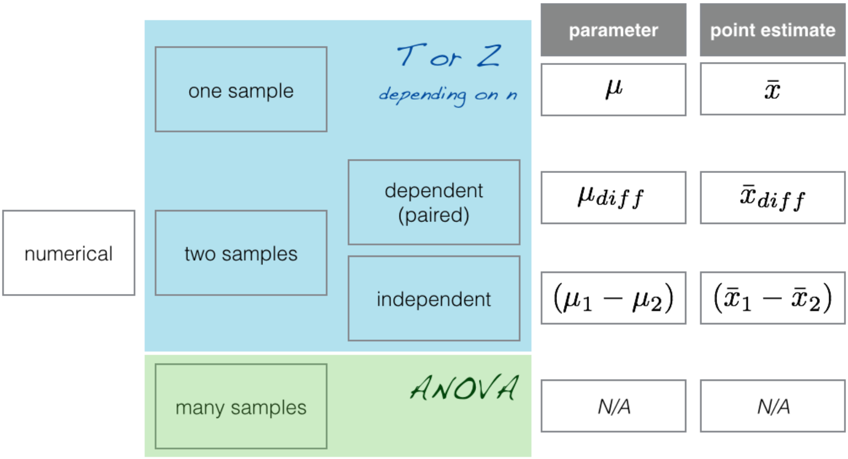
October 21, 2014

# Announcements

- Extra credit on MT
- Project proposal feedback
- PA 4 due Friday at 5pm (extended)

1. Housekeeping
2. Main ideas
   - (1) Roadmap for inference for numerical data
   - (2) ANOVA for comparing many means
   - (3) Within group variability vs. between group variability
   - (4) First ANOVA, then multiple comparisons
3. Application exercise
4. Recap

# (1) Roadmap for inference for numerical data



# (2) ANOVA for comparing many means

Different framework than before: instead of comparing point estimate to null value, compare variability within and between groups.

# (2) ANOVA for comparing many means

Different framework than before: instead of comparing point estimate to null value, compare variability within and between groups.

*$Z/T$ test* Compare means from *two* groups to see whether they are so far apart that the observed difference cannot reasonably be attributed to sampling variability. The test statistic is a ratio.

$$H_0 : \mu_1 = \mu_2 \qquad Z/T = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{SE(\bar{x}_1 - \bar{x}_2)}$$

# (2) ANOVA for comparing many means

Different framework than before: instead of comparing point estimate to null value, compare variability within and between groups.

*Z/T test* Compare means from *two* groups to see whether they are so far apart that the observed difference cannot reasonably be attributed to sampling variability. The test statistic is a ratio.

$$H_0 : \mu_1 = \mu_2 \qquad Z/T = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{SE(\bar{x}_1 - \bar{x}_2)}$$

*ANOVA* Compare the means from *more than two* groups to see whether they are so far apart that the observed differences cannot all reasonably be attributed to sampling variability. The test statistic is a ratio.

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k \qquad F = \frac{\text{variability bet. groups}}{\text{variability w/in groups}}$$

# (3) Within group variability vs. between group variability

Very little variability *within groups*
(observations within a group alike)

+

# (3) Within group variability vs. between group variability

Very little variability *within groups*
(observations within a group alike)

+
lots of variability *between groups*
(groups very different
from each other)
↓

# (3) Within group variability vs. between group variability

Very little variability *within groups*
(observations within a group alike)

+

lots of variability *between groups*
(groups very different
from each other)

↓

likely significant ANOVA

# (3) Within group variability vs. between group variability

Very little variability *within groups*
(observations within a group alike)

+

lots of variability *between groups*
(groups very different
from each other)

↓

likely significant ANOVA

Lots of variability *within groups*
(observations within a group
all over the place)

+

# (3) Within group variability vs. between group variability

Very little variability *within groups*
(observations within a group alike)

+

lots of variability *between groups*
(groups very different
from each other)

↓

likely significant ANOVA

Lots of variability *within groups*
(observations within a group
all over the place)

+

little variability *between groups*
(groups not so different
from each other)

↓

# (3) Within group variability vs. between group variability

Very little variability *within groups*
(observations within a group alike)

+

lots of variability *between groups*
(groups very different from each other)

↓

likely significant ANOVA

Lots of variability *within groups*
(observations within a group all over the place)

+

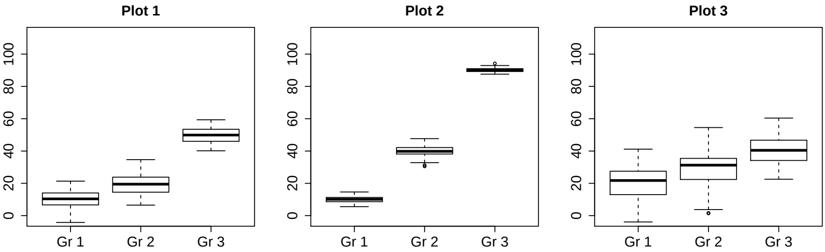
little variability *between groups*
(groups not so different from each other)

↓

likely not significant ANOVA

## Clicker question

Order the plots with respect to *within groups* variability (low to high).




(a) Plot 1, Plot 2, Plot 3

(b) Plot 1, Plot 3, Plot 1

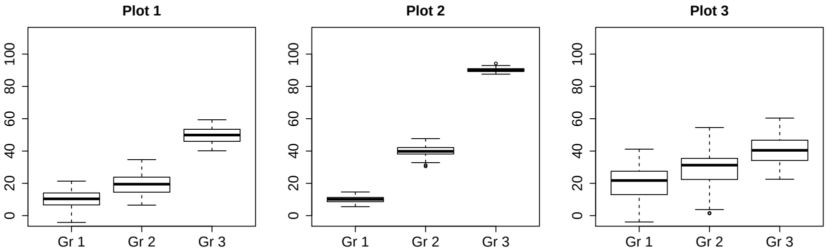
(c) Plot 2, Plot 1, Plot 3

(d) Plot 2, Plot 3, Plot 1

(e) Plot 3, Plot 1, Plot 2

## Clicker question

Order the plots with respect to *within groups* variability (low to high).




(a) Plot 1, Plot 2, Plot 3

(b) Plot 1, Plot 3, Plot 1

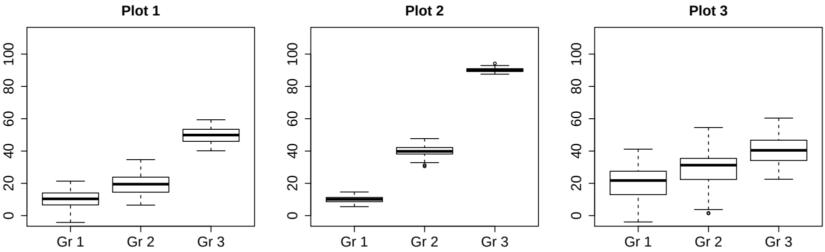
(c) *Plot 2, Plot 1, Plot 3*

(d) Plot 2, Plot 3, Plot 1

(e) Plot 3, Plot 1, Plot 2

## Clicker question

Order the plots with respect to *between groups* variability (low to high).




(a) Plot 1, Plot 2, Plot 3

(b) Plot 1, Plot 3, Plot 1

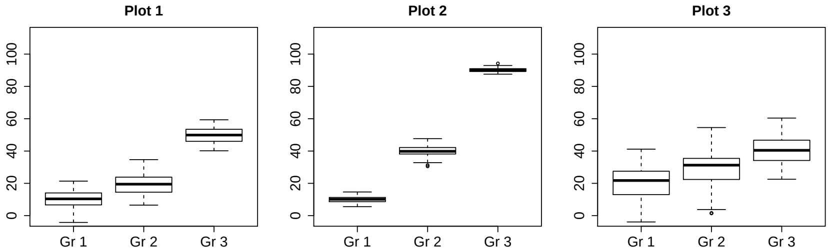
(c) Plot 2, Plot 1, Plot 3

(d) Plot 2, Plot 3, Plot 1

(e) Plot 3, Plot 1, Plot 2

## Clicker question

Order the plots with respect to *between groups* variability (low to high).




(a) Plot 1, Plot 2, Plot 3

(b) Plot 1, Plot 3, Plot 1

(c) Plot 2, Plot 1, Plot 3

(d) Plot 2, Plot 3, Plot 1

(e) *Plot 3, Plot 1, Plot 2*

# (4) First ANOVA, then multiple comparisons

- ANOVA: $H_A$: At least one pair of means are different from each other.

# (4) First ANOVA, then multiple comparisons

- ANOVA: $H_A$: At least one pair of means are different from each other.
- If $H_0$ is rejected, and data provide evidence for $H_A$, we still don't know which groups have differing means.

# (4) First ANOVA, then multiple comparisons

- ANOVA: $H_A$: At least one pair of means are different from each other.
- If $H_0$ is rejected, and data provide evidence for $H_A$, we still don't know which groups have differing means.
- Conduct tests that compare each pair of means to each other: *multiple comparisons*.

# (4) First ANOVA, then multiple comparisons

- ANOVA: $H_A$: At least one pair of means are different from each other.
- If $H_0$ is rejected, and data provide evidence for $H_A$, we still don't know which groups have differing means.
- Conduct tests that compare each pair of means to each other: *multiple comparisons*.
- Each one of these tests might potentially commit a Type 1 error, with a likelihood of $\alpha$.

# (4) First ANOVA, then multiple comparisons

- ANOVA: $H_A$: At least one pair of means are different from each other.
- If $H_0$ is rejected, and data provide evidence for $H_A$, we still don't know which groups have differing means.
- Conduct tests that compare each pair of means to each other: *multiple comparisons*.
- Each one of these tests might potentially commit a Type 1 error, with a likelihood of $\alpha$.
- Therefore, in order to avoid inflating the Type 1 error rate, run each of these tests at a modified (lower) significance level.

$$\alpha^{\star} = \frac{\alpha}{number\ of\ tests}$$

**Application exercise:**
**4.2 ANOVA - Part 1**

See course website for instructions

**Application exercise:**
**4.3 ANOVA - Part 2**

See course website for instructions

# Conditions for ANOVA

1. Independence: random sampling / assignment + 10% rule
2. Approximately normal – can be relaxed if $n$ is large
3. Constant variance – can be relaxed if $n$ is consistent

# Misc. notes on ANOVA

- No such things as a confidence interval for ANOVA, since there is no parameter of interest to estimate.
- You could use a Z test for multiple comparisons if the sample sizes are large enough, but now that you're familiar with the T test (which can be used for small and large samples) you might as well use that for all inference on single means or comparing two means.
- The F distribution is always positive since the F statistic is the ratio of two variabilities (sums of squares). Also, we always shade above the observed F statistic since evidence for the alternative hypothesis means a higher ratio of the between to within variability.