Week 7: Tests for Differences

Session 1

Spring 2020

iClicker Question 1

A `two-sample t-test`

B `Mann-Whitney paried test`

C `Mann-Whitney U test`

D `paired t-test`

Announcements

This week

A question!

Having just analyzed some fish counts data in 16 lakes in Massachusetts, Thorsten found a significant ‘lake’ effect using an ANOVA, i.e., the mean number of fish was not the same in all lakes.

1. What would Thorsten do to find out *which lakes were different from eachother*?

A) A series of t-tests

B) A Tukey Honest Significant Difference test

C) A Kruskal-Wallis test

A question!

I am interested in testing whether there is a significant difference between the population sizes of fish in 30 low salinity lakes and 30 high salinity lakes:

1. Which statistical test should I use?

A) A t-test
B) A One-Way ANOVA
C) A Chi-square test
D) A Two-Way ANOVA

A question!

I am interested in testing whether there is a significant difference between the population sizes of fish in 30 low salinity lakes and 30 high salinity lakes:

2. Which is the test statistic for the test?

A) t
B) F
C) r
D) χ^2

A question!

I am interested in testing whether there is a significant difference between the population sizes of fish in 30 low salinity lakes and 30 high salinity lakes. In fact, I actually sampled 10 large, 10 medium, and 10 small lakes in each of the high and low salinity lakes. I want to explore whether there are differences in population size based on lake salinity and lake size.

3. Now which statistical test should I use?

A) A t-test
B) A One-Way ANOVA
C) A Chi-square test
D) A Two-Way ANOVA

A question!

I am interested in testing whether there is a significant difference between the population sizes of fish in 30 low salinity lakes and 30 high salinity lakes. In fact, I actually sampled 10 large, 10 medium, and 10 small lakes in each of the high and low salinity lakes. I want to explore whether there are differences in population size based on lake salinity and lake size.

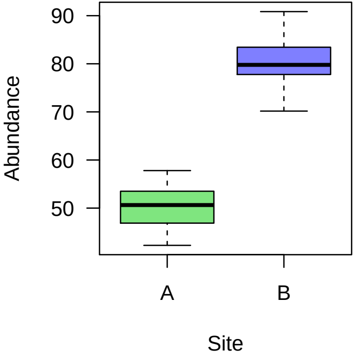
4. Now Which is the test statistic for the test?

A) t
B) F
C) r
D) χ^2

Comparing differences - two samples

Two samples:

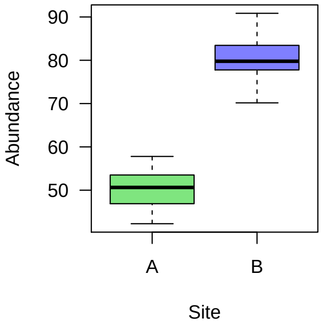
- Which test do we use?



Comparing differences - two samples

Two samples:

- the t-test?
- test whether group means differ significantly
- H_0: there is no significant difference between the means
- H_1: there is a significant difference between the means



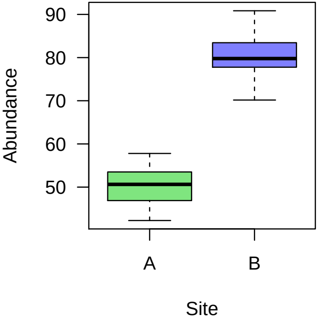
Comparing differences - two smaples

Two samples:

- the t-test?
- test whether group means differ significantly
- H_0: there is no significant difference between the means
- H_1: there is a significant difference between the means

Significance based on:

- t-statistic: $t = \frac{|\bar{x}_a - \bar{x}_b|}{\sqrt{\frac{s_a^2}{n_a} + \frac{s_b^2}{n_b}}}$
- degrees of freedom
- p-value



Comparing differences - more than two samples

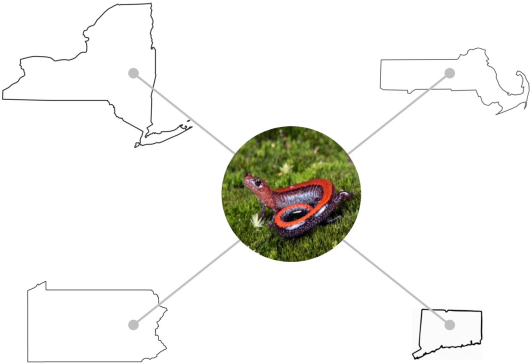
What about if there are more than 2 samples?

- can you think of any examples?

Comparing multiple groups - examples

Regional differences in salamander abundance:

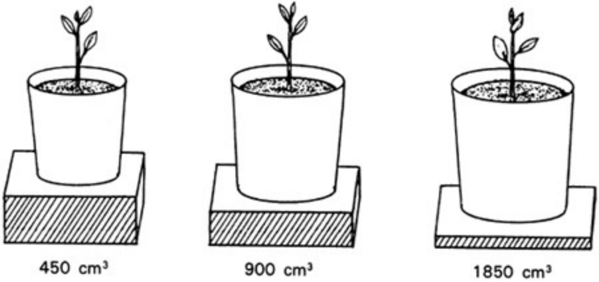
- comparing multiple populations
- quantify the differences between populations



Comparing multiple groups - examples

Plant growth related to available resources (pot size):

- comparing multiple treatments
- quantify the effects of resource availability



Comparing multiple groups - examples

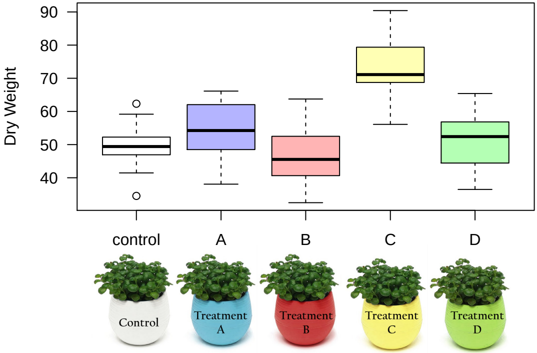
Plants productivity (dry mass in grams) related to fertilizer treatment

- do our treatments influence biomass production?
- is there a positive effect relative to a control?

Comparing multiple groups - examples

When there are more than 2 groups

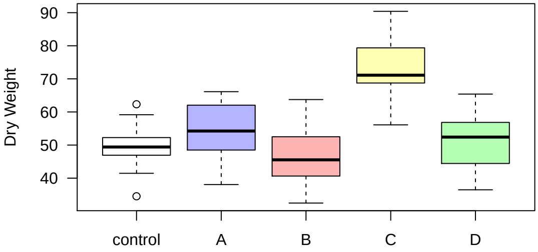
- t-test doesn't help
  - need to do all possible pairs
  - time consuming
  - get spurious differences just by chance



Comparing multiple groups - ANOVA

Analysis of Variance (ANOVA):

- statistical test for testing for differences among >2 groups
- ANOVA and t-test are identical when there are 2 groups
- one factor/group/category (*One-way ANOVA*)



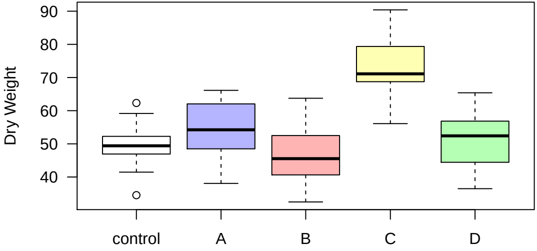
Comparing multiple groups - ANOVA

Analysis of Variance (ANOVA):

- statistical test for testing for differences among >2 groups
- ANOVA and t-test are identical when there are 2 groups
- one factor/group/category (*One-way ANOVA*)

Assumption:

- data are normally distributed



Comparing multiple groups - ANOVA

Analysis of Variance (ANOVA):

- statistical test for testing for differences among >2 groups
- ANOVA and t-test are identical when there are 2 groups
- one factor/group/category (*One-way ANOVA*)

Assumption:

- data are normally distributed

Hypotheses:

?

Comparing multiple groups - ANOVA

Analysis of Variance (ANOVA):

- statistical test for testing for differences among >2 groups
- ANOVA and t-test are identical when there are 2 groups
- one factor/group/category (*One-way ANOVA*)

Assumption:

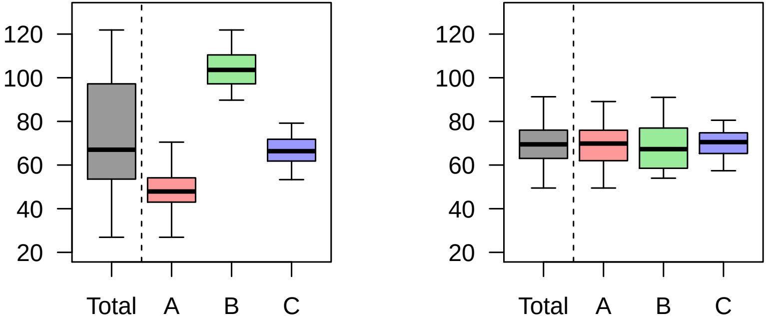
- data are normally distributed

Hypotheses:

- H_0: there are no significant differences between the means
  - all means are equal
- H_1: there are significant differences between the means
  - all means are not equal

ANOVA explained

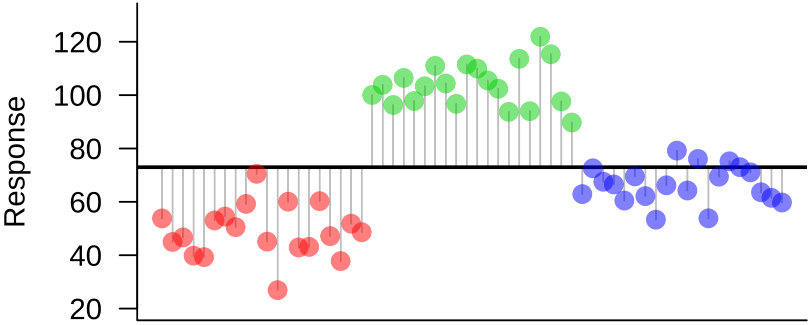
The ANOVA partitions the *total* variation into *within* sample variation with *between* sample variation to determine whether samples come from a single distribution or not.



ANOVA and the Sums of Squares

- *Total* sums of squares (SS_T)

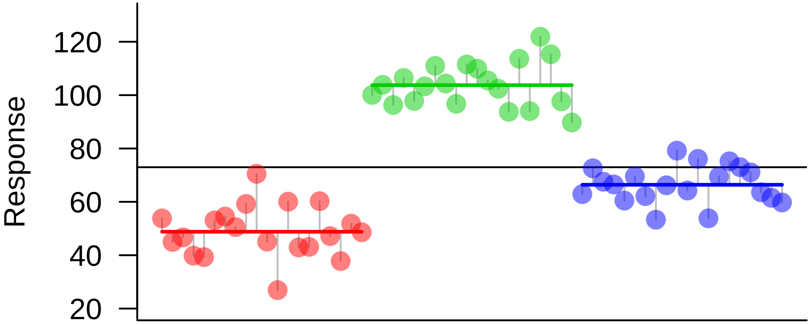
$$SS_T = \sum (x - \bar{x})^2$$



ANOVA and the Sums of Squares

- *Within-sample* sums of squares (SS_T)
- add up the within sample SS

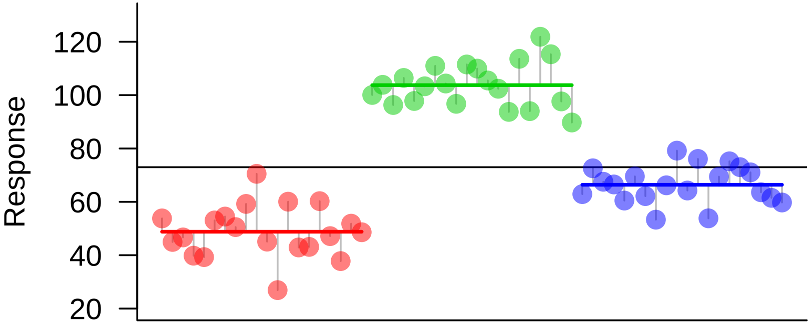
$$SS_W = \sum(x_1 - \bar{x}_1)^2 + \sum(x_2 - \bar{x}_2)^2 + \sum(x_3 - \bar{x}_3)^2$$



ANOVA and the Sums of Squares

- *Within-sample* sums of squares (SS_T)
- more generally (g is the number of groups)

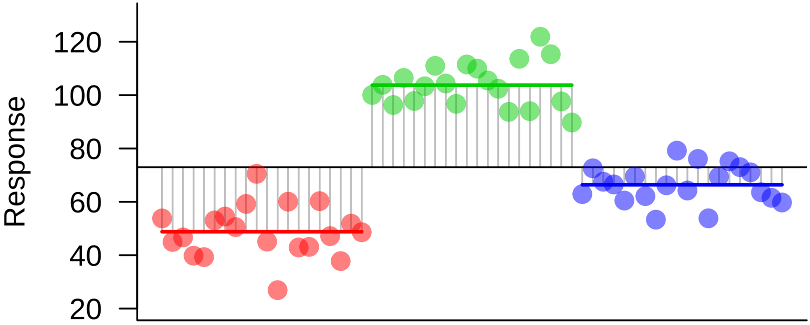
$$SS_W = \sum_g \sum_i (x_{ig} - \bar{x}_g)^2$$



ANOVA and the Sums of Squares

- *Between-sample* sums of squares (SS_T)
- add up the differences in the means

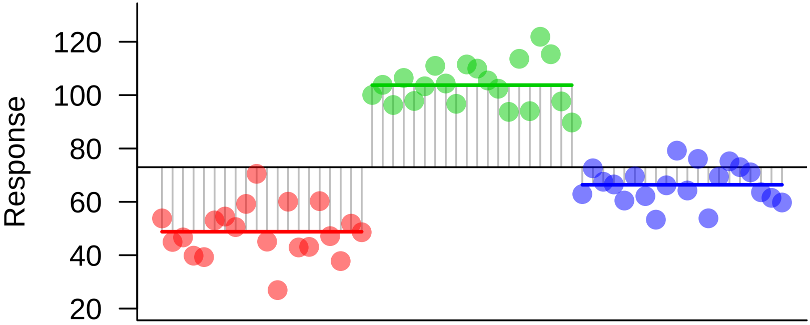
$$SS_B = n_1(\bar{x}_1 - \bar{x})^2 + n_2(\bar{x}_2 - \bar{x})^2 + n_3(\bar{x}_3 - \bar{x})^2$$



ANOVA and the Sums of Squares

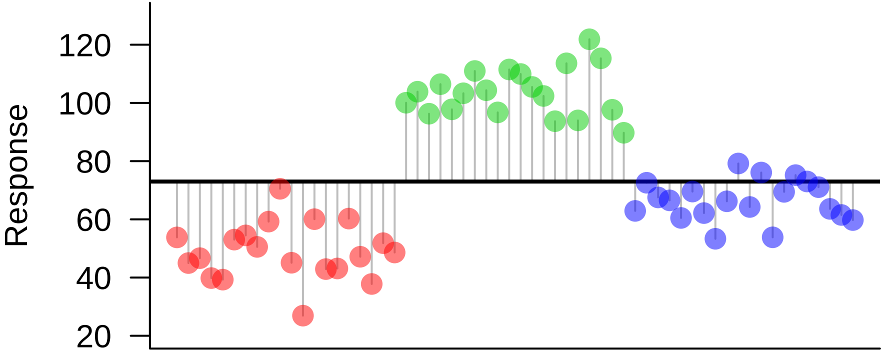
- *Between-sample* sum of squares (SS_T)
- more generally (g is the number of groups)

$$SS_B = \sum_g n_g(\bar{x}_g - \bar{x})^2$$



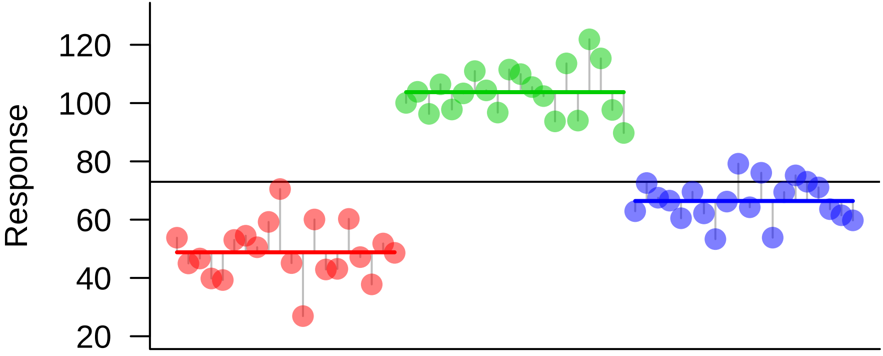
ANOVA and the Sums of Squares

Total:



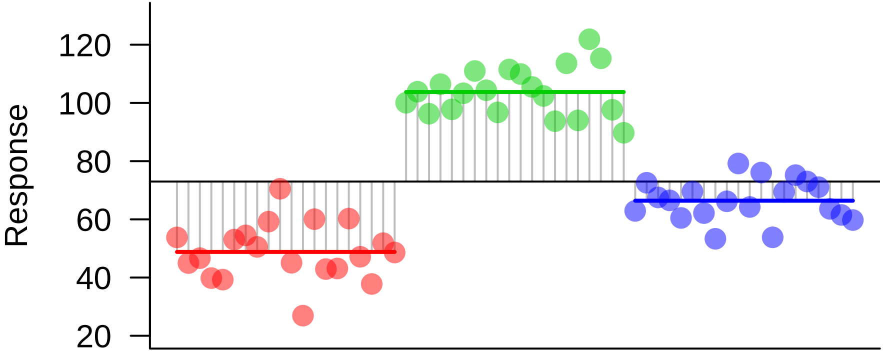
ANOVA and the Sums of Squares

Within group:



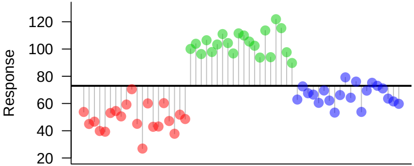
ANOVA and the Sums of Squares

Between group:

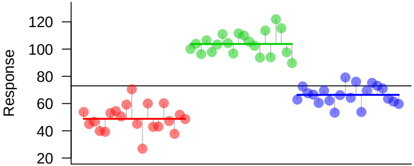


ANOVA and the Sums of Squares

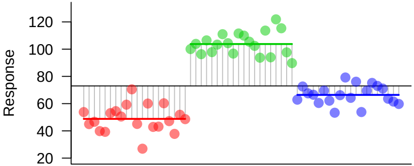
$$SS_T = \sum (x - \bar{x})^2$$



$$SS_W = \sum_g \sum_i (x_{ig} - \bar{x}_g)^2$$



$$SS_B = \sum_g n_g (\bar{x}_g - \bar{x})^2$$



ANOVA degrees of freedom

If we define the following:

- n is the total sample size (number of observations)
- g is the number of groups/samples

ANOVA degrees of freedom

If we define the following:

- n is the total sample size (number of observations)
- g is the number of groups/samples

Then the degrees of freedom (df) are:

- Total: $df_T = n - 1$
- Within: $df_W = g - 1$
- Between: $df_B = n - g$

ANOVA the *mean square*

The mean square (MS) is the sum of squares divided by the degrees of freedom:

$$MS = SS/df$$

So:

- Total: $MS_T = SS_T/df_T$
- Within: $MS_W = SS_W/df_W$
- Between: $MS_B = SS_B/df_B$

ANOVA all the ingredients

	SS	df	MS
Total	$\sum(x-\bar{x})^2$	$n-1$	SS_T/df_T
Within	$\sum_g \sum_i (x_{ig}-\bar{x}_j)^2$	$n-g$	SS_W/df_W
Between	$\sum_g n_g(\bar{x}_g-\bar{x})^2$	$g-1$	SS_B/df_B

ANOVA the statistical test

ANOVA results are usually presented in an ANOVA table

\begin{center}

Source of variation	SS	df	MS	F	p
Between	SS_B	df_B	MS_B		
Within	SS_W	df_W	MS_W		
Total	SS_T	df_T	–		

\end{center}

ANOVA the statistical test

ANOVA results are usually presented in an ANOVA table

\begin{center}

Source of variation	SS	df	MS	F	p
Between	SS_B	df_B	MS_B		
Within	SS_W	df_W	MS_W		
Total	SS_T	df_T	–		

\end{center}

- F is the test statistic for the ANOVA

$$F = \frac{MS_B}{MS_W}$$

ANOVA the statistical test

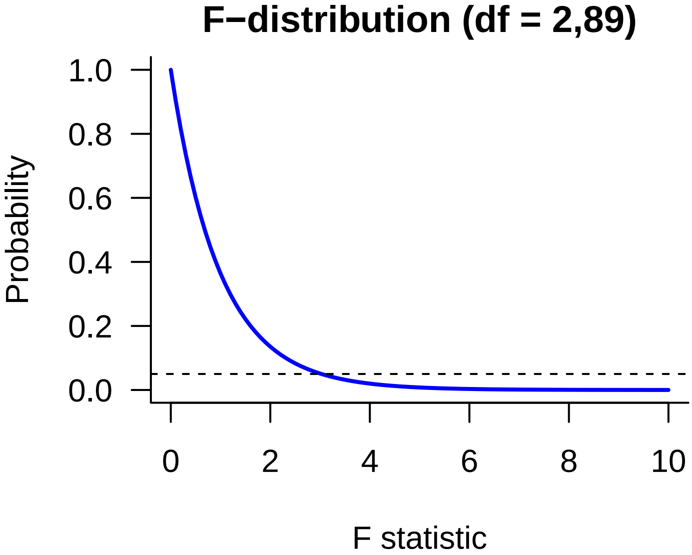
ANOVA results are usually presented in an ANOVA table

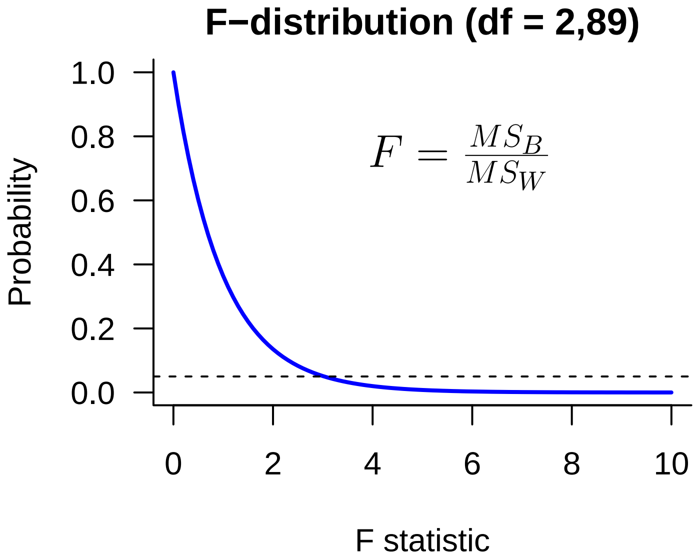
\begin{center}

Source of variation	SS	df	MS	F	p
Between	SS_B	df_B	MS_B		
Within	SS_W	df_W	MS_W		
Total	SS_T	df_T	–		

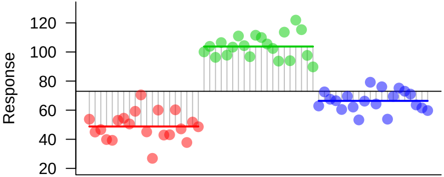
\end{center}

- p is the probability of observing the F statistic with a given degrees of freedom if the null hypothesis is true:
  - null hypothesis is 'no difference between the means'
  - based on the F-distribution

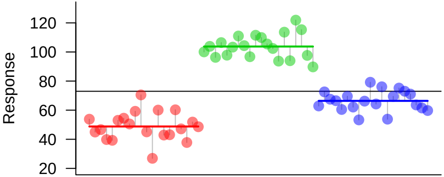

F−distribution (df = 2,89)
Probability
1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
8
10
F statistic


F−distribution (df = 2,89)
$F = \frac{MS_B}{MS_W}$
Probability
1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
8
10
F statistic

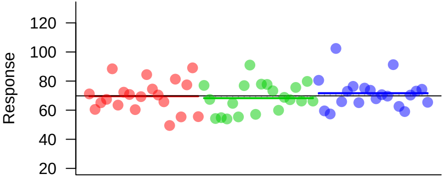
ANOVA and the Sums of Squares



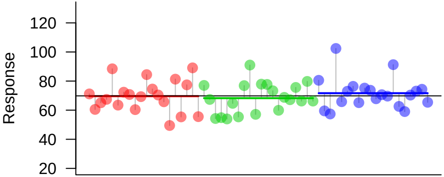
$$F = \frac{MS_B}{MS_W}$$



ANOVA and the Sums of Squares



$$F = \frac{MS_B}{MS_W}$$



ANOVA the p value

Hypotheses:

- H_0: there are no significant differences between the means
  - all means are equal
- H_1: there are significant differences between the means
  - all means are not equal

When do we reject or fail to reject the null hypothesis?

ANOVA the p value

Hypotheses:

- H_0: there are no significant differences between the means
  - all means are equal
- H_1: there are significant differences between the means
  - all means are not equal

When do we reject or fail to reject the null hypothesis?

- if F is large, then p is small
- if $p < 0.05$ we reject the null hypothesis
- if $p > 0.05$ we *fail to* reject the null hypothesis

Pairwise comparisons with ANOVA

The F statistic tells us whether there are differences, but *not* what the differences are:

- *Cannot* use t-tests to make pairwise comparisons
  - multiple t-tests will lead to significant results by chance

Pairwise comparisons with ANOVA

The F statistic tells us whether there are differences, but *not* what the differences are:

- Instead we conduct *Post-hoc* testing
  - Tukey Honest Significant Difference test (Tukey HSD)
  - accounts for multiple tests being conducted
  - calculation of a t-statistic
  - a pair, so degrees of freedom is 1
  - 5% critical value for $df = 1$ is 4.303
    - if $t > 4.303$ then $p < 0.05$

Pairwise comparisons with ANOVA

The F statistic tells us whether there are differences, but *not* what the differences are:

- Instead we conduct *Post-hoc* testing
  - Tukey Honest Significant Difference test (Tukey HSD)
  - accounts for multiple tests being conducted
  - calculation of a t-statistic
  - a pair, so degrees of freedom is 1
  - 5% critical value for $df = 1$ is 4.303
    - if $t > 4.303$ then $p < 0.05$

$$t_{a,b} = \frac{|\bar{x}_a - \bar{x}_b|}{\sqrt{\frac{MS_W\left(\frac{1}{n_a}+\frac{1}{n_b}\right)}{2}}}$$

Pairwise comparisons with ANOVA

The F statistic tells us whether there are differences, but *not* what the differences are:

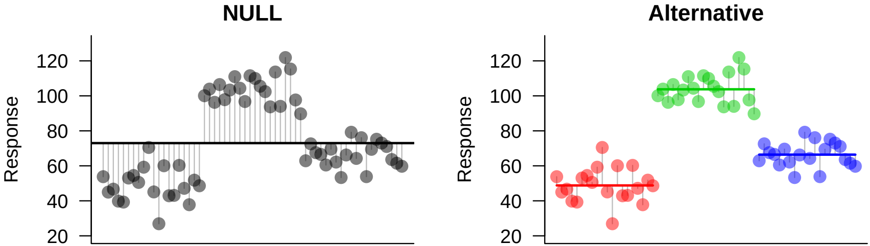
- Instead we conduct *Post-hoc* testing
  - Tukey Honest Significant Difference test (Tukey HSD)
  - accounts for multiple tests being conducted
  - calculation of a t-statistic
  - a pair, so degrees of freedom is 1
  - 5% critical value for $df = 1$ is 4.303
    - if $t > 4.303$ then $p < 0.05$

	A	B	C
A	-	$t_{A,B}$	$t_{A,C}$
B	-	-	$t_{B,C}$
C	-	-	-

ANOVA Recap

Comparing differences between >2 samples (groups) using ANOVA

- null hypothesis:
  - no difference between the samples
  - data are from the same population
- alternative hypothesis:
  - sample means are different
  - data from the different populations



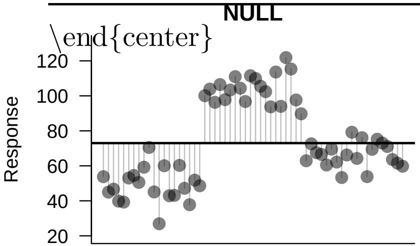
ANOVA Recap

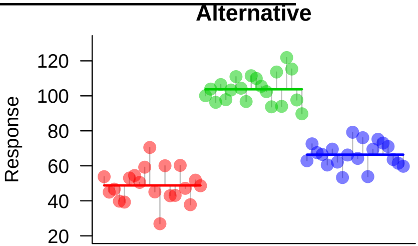
Comparing differences between >2 groups using ANOVA

\begin{center}

Source of variation	SS	df	MS	F	p
Between	SS_B	df_B	MS_B	$\frac{MS_B}{MS_W}$	
Within	SS_W	df_W	MS_W		
Total	SS_T	df_T	$-$		

\end{center}

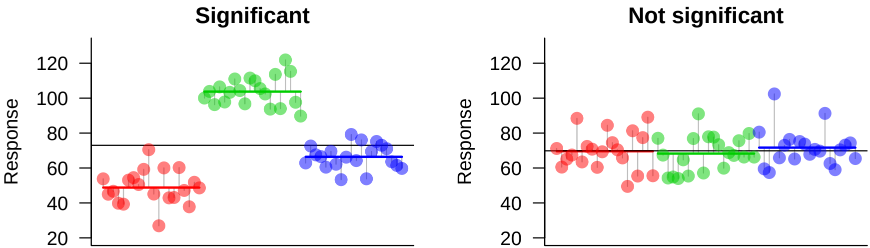




ANOVA Recap

Comparing differences between >2 groups using ANOVA

- Essentially comes down to:
  - a model with one mean *or* a model with a mean per group
  - which model best explains the data
  - which model significantly reduces the sums of squares



More than one factor with ANOVA

So far we have looked at multiple levels within a single factor

- factor: a single categorical predictor variable
- level: the categories within a factor

In some cases, we may be interested in >1 factor

- 2 factors: *two-way* ANOVA
- 3 factors: *three-way* ANOVA
- $\cdots$ multi-way ANOVA

Two-way ANOVA

Let's use a grazing example:

\begin{center}

	Site	
Grazing Treatment	Top	Lower
Lo	9	7
Lo	11	6
Lo	6	5
Mid	14	14
Mid	17	17
Mid	19	15
Hi	28	44
Hi	31	38
Hi	32	37

\end{center}

Two-way ANOVA

Lets use the example from the book (in R looks like this):

```
graze
```

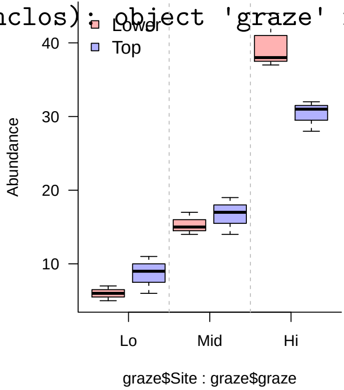
```
## Error in eval(expr, envir, enclos): object 'graze' not f
```

Two-way ANOVA

Lets use the example from the book (in `R` looks like this):

```
graze
```

```
## Error in eval(expr, envir, enclos): object 'graze' not f
```

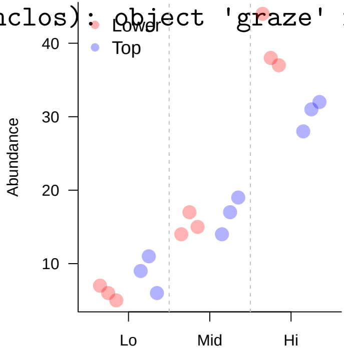


Two-way ANOVA

Lets use the example from the book (in `R` looks like this):

```
graze
```

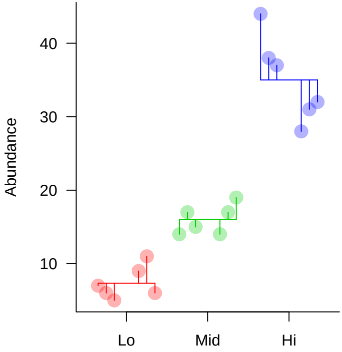
```
## Error in eval(expr, envir, enclos): object 'graze' not f
```



Conducting the ANOVA

Step one:

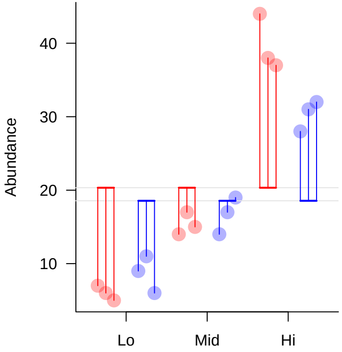
- SS for each factor
  - graze
  - site
- $SS_{graze} = \sum(x_{i,graze} - \bar{x}_{graze})^2$
- Ignore site grouping



Conducting the ANOVA

Step one:

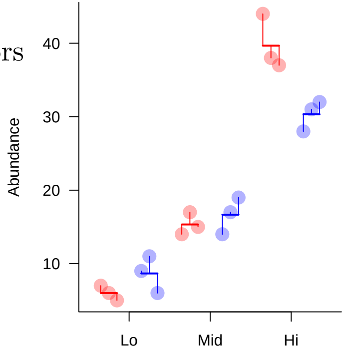
- SS for each factor
  - graze
  - site
- $SS_{site} = \sum(x_{i,site} - \bar{x}_{site})^2$
- Ignore graze grouping



Conducting the ANOVA

Step two:

- SS for each combinations of factors
- Treat all groupings as unique
- $SS_{within} = (x_{i,g} - \bar{x}_g)^2$



Conducting the ANOVA

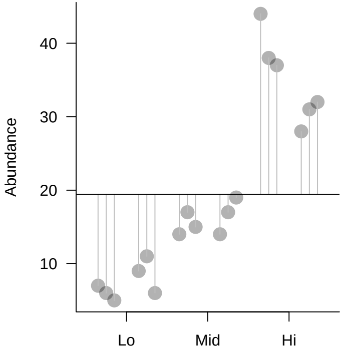
Step three:

- Sums of squares of both factors
- $SS_{both} = SS_{total} - SS_{graze} - SSsite - SS_{within}$

Conducting the ANOVA

Step four:

- Total sums of squares
- $SS_{total} = \sum(x_i - \bar{x})^2$
- the *null* model
- Ignore all group structure



Conducting the ANOVA - sums of squares

\begin{center}

	SS	df	MS	F	p
Graze	SS_{graze}				
Site	SS_{site}				
Both factors(interaction)	SS_{both}				
Within group	SS_{within}				
Total	SS_{total}				

\end{center}

Degrees of freedom

In general:

- Factor 1 (F1): number of levels - 1
- Factor 2 (F2): number of levels - 1
- Within: n - (levels in F1 $\times$ levels in F2)
- Total: n - 1

Degrees of freedom

In general:

- Factor 1 (F1): number of levels - 1
- Factor 2 (F2): number of levels - 1
- Within: n - (levels in F1 $\times$ levels in F2)
- Total: n - 1

Grazing example:

- Graze: $3 - 1 = 2$
- Site: $2 - 1 = 1$
- Within: $18 - (3 \times 2) = 12$
- Total: $18 - 1 = 17$

Degrees of freedom

In general:

- Factor 1 (F1): number of levels - 1
- Factor 2 (F2): number of levels - 1
- Within: n - (levels in F1 $\times$ levels in F2)
- Total: n - 1

\begin{center}

	SS	df	MS	F	p
Graze	SS_{graze}	df_{graze}			
Site	SS_{site}	df_{site}			
Both factors(interaction)	SS_{both}	df_{both}			
Within group	SS_{within}	df_{within}			
Total	SS_{total}	df_{total}			

\end{center}

Mean squares

- the mean squares are calculated by dividing the sums of squares by the degrees of freedom for each element

\begin{center}

	SS	df	MS	F	p
Graze	SS_{graze}	df_{graze}	$MS_{graze} = \frac{SS_{graze}}{df_{graze}}$		
Site	SS_{site}	df_{site}	$MS_{site} = \frac{SS_{site}}{df_{site}}$		
Both factors	SS_{both}	df_{both}	$MS_{both} = \frac{SS_{both}}{df_{both}}$		
Within group	SS_{within}	df_{within}	$MS_{within} = \frac{SS_{within}}{df_{within}}$		
Total	SS_{total}	df_{total}			

\end{center}

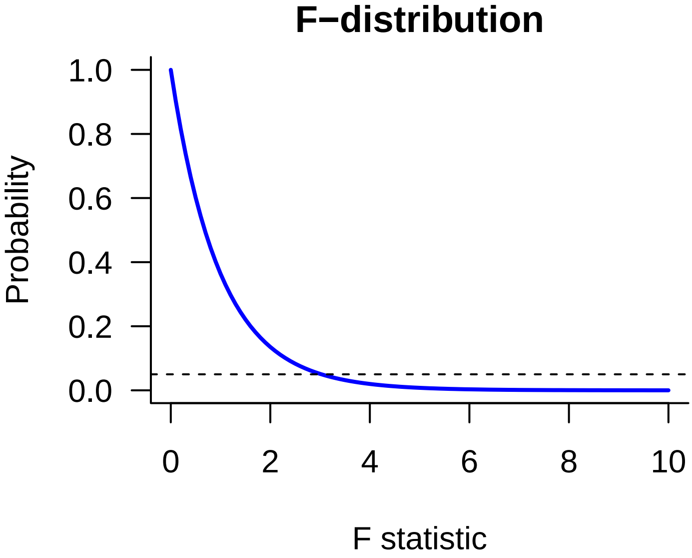
F statistic

- the F-statistic is calculated by taking the element of interest divided by the within group MS (the *error* term)

\begin{center}

	SS	df	MS	F	p
Graze	SS_{graze}	df_{graze}	$MS_{graze} = \frac{SS_{graze}}{df_{graze}}$	$\frac{MS_{graze}}{MS_{within}}$	
Site	SS_{site}	df_{site}	$MS_{site} = \frac{SS_{site}}{df_{site}}$	$\frac{MS_{site}}{MS_{within}}$	
Both factors	SS_{both}	df_{both}	$MS_{both} = \frac{SS_{both}}{df_{both}}$	$\frac{MS_{both}}{MS_{within}}$	
Within group	SS_{within}	df_{within}	$MS_{within} = \frac{SS_{within}}{df_{within}}$		
Total	SS_{total}	df_{total}			

\end{center}


F−distribution
Probability
1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
8
10
F statistic

ANOVA in practice - R

- Read in the data as a data frame

```
graze
```

```
## Error in eval(expr, envir, enclos): object 'graze' not f
```

Any ANOVA in practice - R

- Conduct *any* test using formula syntax

```
oneway.site <- aov(Abundance ~ Site, data = graze)
```

```
## Error in terms.formula(formula, "Error", data = data): o
```

```
summary(oneway.site)
```

```
## Error in summary(oneway.site): object 'oneway.site' not
```

Any ANOVA in practice - R

- Conduct *any* test using formula syntax

```
oneway.site <- aov(Abundance ~ Site, data = graze)
```

```
## Error in terms.formula(formula, "Error", data = data): o
```

```
summary(oneway.site)
```

```
## Error in summary(oneway.site): object 'oneway.site' not
```

Any ANOVA in practice - R

- Conduct *any* test using formula syntax

```
oneway.graze <- aov(Abundance ~ graze, data = graze)
```

```
## Error in terms.formula(formula, "Error", data = data): c
```

```
summary(oneway.graze)
```

```
## Error in summary(oneway.graze): object 'oneway.graze' no
```

Any ANOVA in practice - R

- Conduct *any* test using formula syntax

```
oneway.graze <- aov(Abundance ~ graze, data = graze)
```

```
## Error in terms.formula(formula, "Error", data = data): c
```

```
summary(oneway.graze)
```

```
## Error in summary(oneway.graze): object 'oneway.graze' no
```

Any ANOVA in practice - R

- Conduct *any* test using formula syntax

```
twoway.additive <- aov(Abundance ~ Site + graze, data = gra
```

```
## Error in terms.formula(formula, "Error", data = data): c
```

```
summary(twoway.additive)
```

```
## Error in summary(twoway.additive): object 'twoway.additi
```

Any ANOVA in practice - R

- Conduct *any* test using formula syntax

```
twoway.additive <- aov(Abundance ~ Site + graze, data = gra
```

```
## Error in terms.formula(formula, "Error", data = data): c
```

```
summary(twoway.additive)
```

```
## Error in summary(twoway.additive): object 'twoway.additi
```

Any ANOVA in practice - R

- Conduct *any* test using formula syntax

```
twoway.interaction <- aov(Abundance ~ Site * graze, data =
```

```
## Error in terms.formula(formula, "Error", data = data): o
```

```
summary(twoway.interaction)
```

```
## Error in summary(twoway.interaction): object 'twoway.int
```

Any ANOVA in practice - R

- Conduct *any* test using formula syntax

```
twoway.interaction <- aov(Abundance ~ Site * graze, data =
```

```
## Error in terms.formula(formula, "Error", data = data): o
```

```
summary(twoway.interaction)
```

```
## Error in summary(twoway.interaction): object 'twoway.int
```

Group Exercise - *salamANOVA*

We will conduct three analyses using the *salamANOVA*. We are interested in whether salamander snout-to-vent length (SVL) varies by sex and/or site. The data look like this:

```
str(sals)
```

```
## Error in str(sals): object 'sals' not found
```

- `Site`: there are four sites (`P1A`, `P1B`, `P2A`, `P2B`)
- `Sex`: `M` (male) and `F` (female)
- `SVL`: the snout-to-vent length in mm

Group Exercise - *salamANOVA*

Analysis 1: Does SVL vary by sex?

- What is the null hypothesis?
- Make a plot to visualize the hypothesis.
- What statistical test will you use to test H_0?
- What is the:
  - test statistic for this particulat test (e.g., t, F, etc)
  - degrees of freedom (calculate this)
  - significance level
- Conduct the analysis:
  - what is the value of the test statistic
  - what the p-value
- Write a short paragraph reporting the conclusion, use values from the statistical test to suppo, supported by the results from the test.

Group Exercise - *salamANOVA*

Analysis 2: Does SVL vary by site?

- What is the null hypothesis?
- Make a plot to visualize the hypothesis.
- What statistical test will you use to test H_0?
- What is the:
  - test statistic for this particulat test (e.g., t, F, etc)
  - degrees of freedom (calculate this)
  - significance level
- Conduct the analysis:
  - what is the value of the test statistic
  - what the p-value
- Write a short paragraph reporting the conclusion, use values from the statistical test to suppo, supported by the results from the test.

Group Exercise - *salamANOVA*

Analysis 3: Does SVL vary by sex and/or site?

- What is the null hypothesis?
- Make a plot to visualize the hypothesis.
- What statistical test will you use to test H_0?
- What is the:
  - test statistic for this particulat test (e.g., t, F, etc)
  - degrees of freedom (calculate this)
  - significance level
- Conduct the analysis:
  - what is the value of the test statistic
  - what the p-value
- Write a short paragraph reporting the conclusion, use values from the statistical test to suppo, supported by the results from the test.

Group Exercise - *salamANOVA*

Assignment: Statistical analysis of variation in salamnder SVL.

- Write a report with four sections:
  1. Analysis 1
  2. Analysis 2
  3. Analysis 3
  4. Reflection: how does analysis 3 compare to analyses 1 and 2?
- Sections 1 to 3 sould report on each of the prompts in the previous slides.
- Section 4 is an opportunity to demonstrate your undertanding of the material covered over the previous weeks.
- Assignment due: 11.55pm Tuesday November $20^{t}h$