Statistical Methods for Plant Biology

PBIO 3150/5150

Anirudh V. S. Ruhil

April 5, 2016

The Voinovich School of Leadership and Public Affairs

Table of Contents

Factorial Designs

Two-Way ANOVA

- We may have more than one factor we want to consider
- Let us assume we have TWO factors (i.e., independent variables)
- For example, we are interested in looking at how temperature and humidity in a room influence student performance on a test. That is, we have an outcome (Math Score) and two factors (aka independent variables):
  1. Factor A: Humidity – two levels (Low vs. High)
  2. Factor B: Temperature – three levels (70/80/90 degrees F)
- How might we test whether (a) Temperature and/or Humidity influence Math performance, and (b) if the effects of one independent variable (for e.g., Temperature) are constant across the values of the other independent variable (for e.g., Humidity)?

Simulated Data

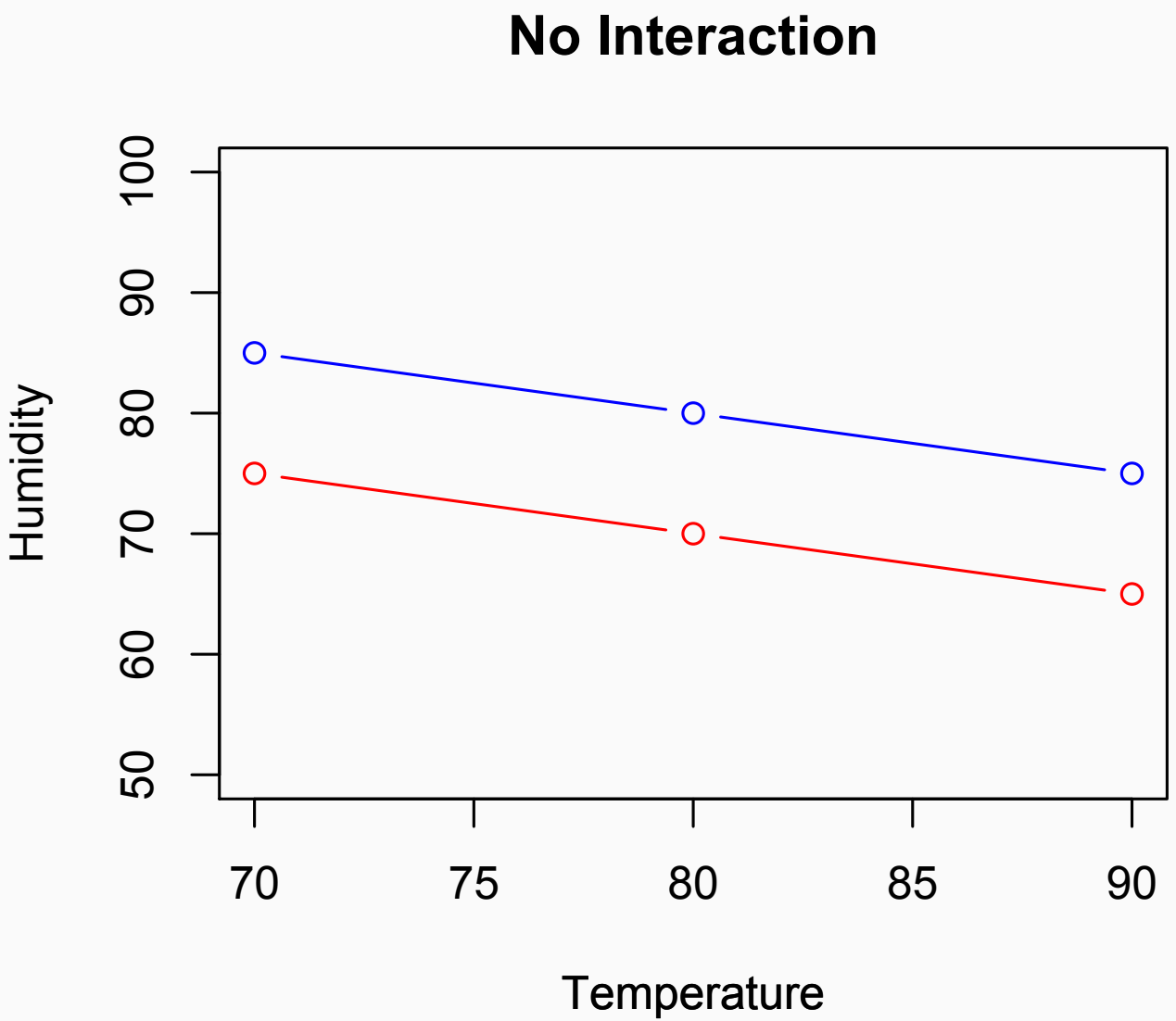
		Factor B: Room Temperature 70^0	80^0	90^0	
Factor A: Humidity	Low	$\bar{y} = 85$	$\bar{y} = 80$	$\bar{y} = 75$	$\bar{y}_{low} = 80$
	High	$\bar{y} = 75$	$\bar{y} = 70$	$\bar{y} = 65$	$\bar{y}_{high} = 70$
		$\bar{y}_{70} = 80$	$\bar{y}_{80} = 75$	$\bar{y}_{90} = 70$	

- Main Effect of Factor A (Humidity): Difference between means for high and low humidity $= 80 - 70 = 10$
- Main Effect for Factor B (Temperature): Difference between means for 70^0, 80^0, and 90^0 temperature $= 80 - 75 = 5; 75 - 70 = 5$
- Null Hypothesis for testing effects of Factor A is $H_0 : \mu_{A1} = \mu_{A2}$
- Null Hypothesis for testing effects of Factor B is $H_0 : \mu_{B1} = \mu_{B2} = \mu_{B3}$

No Interaction Effects

		Factor B: Room Temperature 70^0	80^0	90^0	
Factor A: Humidity	Low	$\bar{y} = 85$	$\bar{y} = 80$	$\bar{y} = 75$	$\bar{y}_{low} = 80$
	High	$\bar{y} = 75$	$\bar{y} = 70$	$\bar{y} = 65$	$\bar{y}_{high} = 70$
		$\bar{y}_{70} = 80$	$\bar{y}_{80} = 75$	$\bar{y}_{90} = 70$	

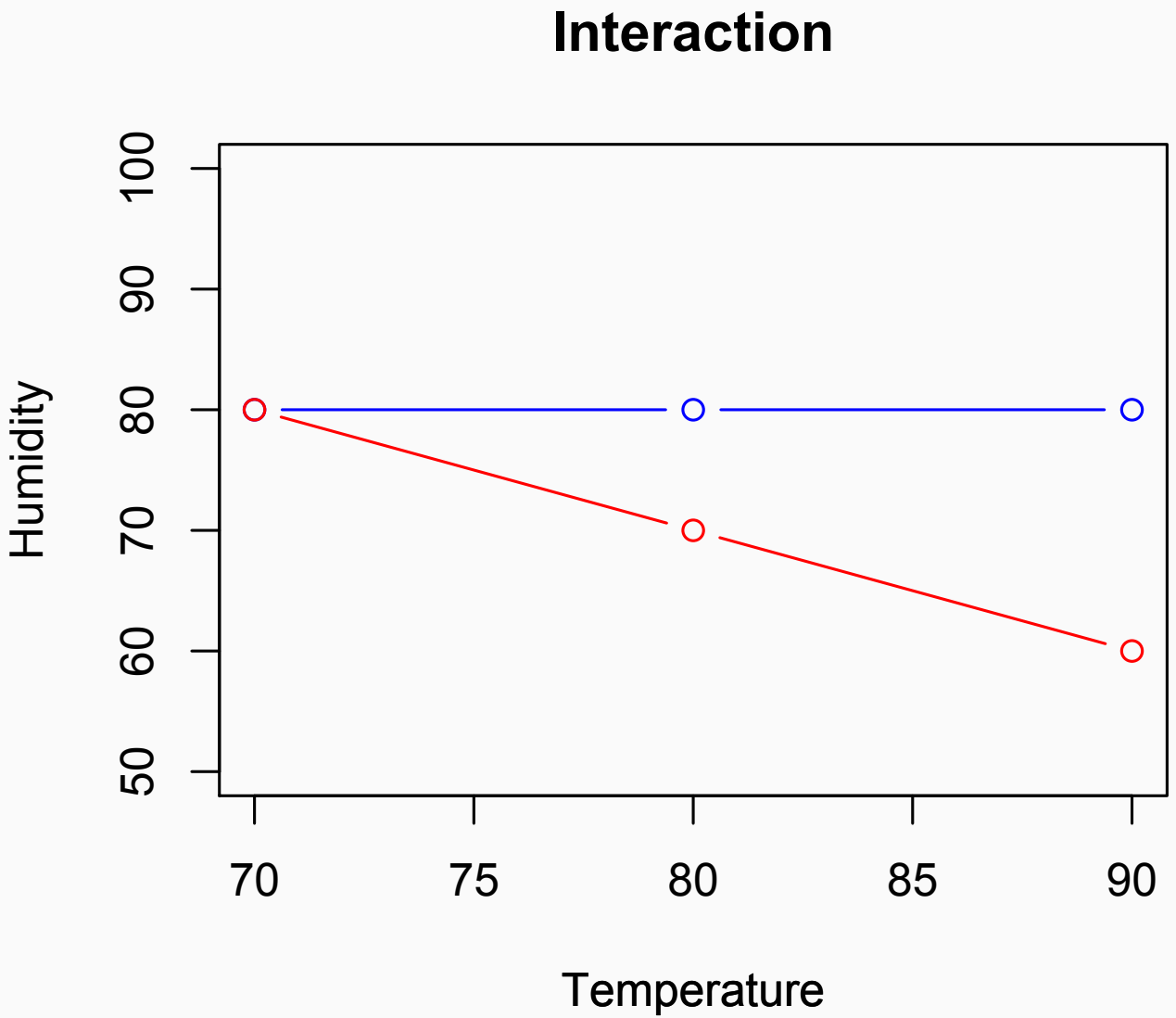
- Note how $\bar{y}$ drops as Temperature increases
- Note how $\bar{y}$ drops as Humidity rises
- Note that at 70^0 there is a difference of 10 between Low/High humidity
- Note that at 80^0 there is a difference of 10 between Low/High humidity
- Note that at 90^0 there is a difference of 10 between Low/High humidity
- Note that at Low Humidity there is a difference of 5 for every 10^0 rise in temperature
- Note that at High Humidity there is a difference of 5 for every 10^0 rise in temperature


No Interaction
Humidity
100
90
80
70
60
50
70
75
80
85
90
Temperature

Interaction Effects

		Factor B: Room Temperature 70^0	80^0	90^0	
Factor A: Humidity	Low	$\bar{y} = 80$	$\bar{y} = 80$	$\bar{y} = 80$	$\bar{y}_{low} = 80$
	High	$\bar{y} = 80$	$\bar{y} = 70$	$\bar{y} = 60$	$\bar{y}_{high} = 70$
		$\bar{y}_{70} = 80$	$\bar{y}_{80} = 75$	$\bar{y}_{90} = 70$	

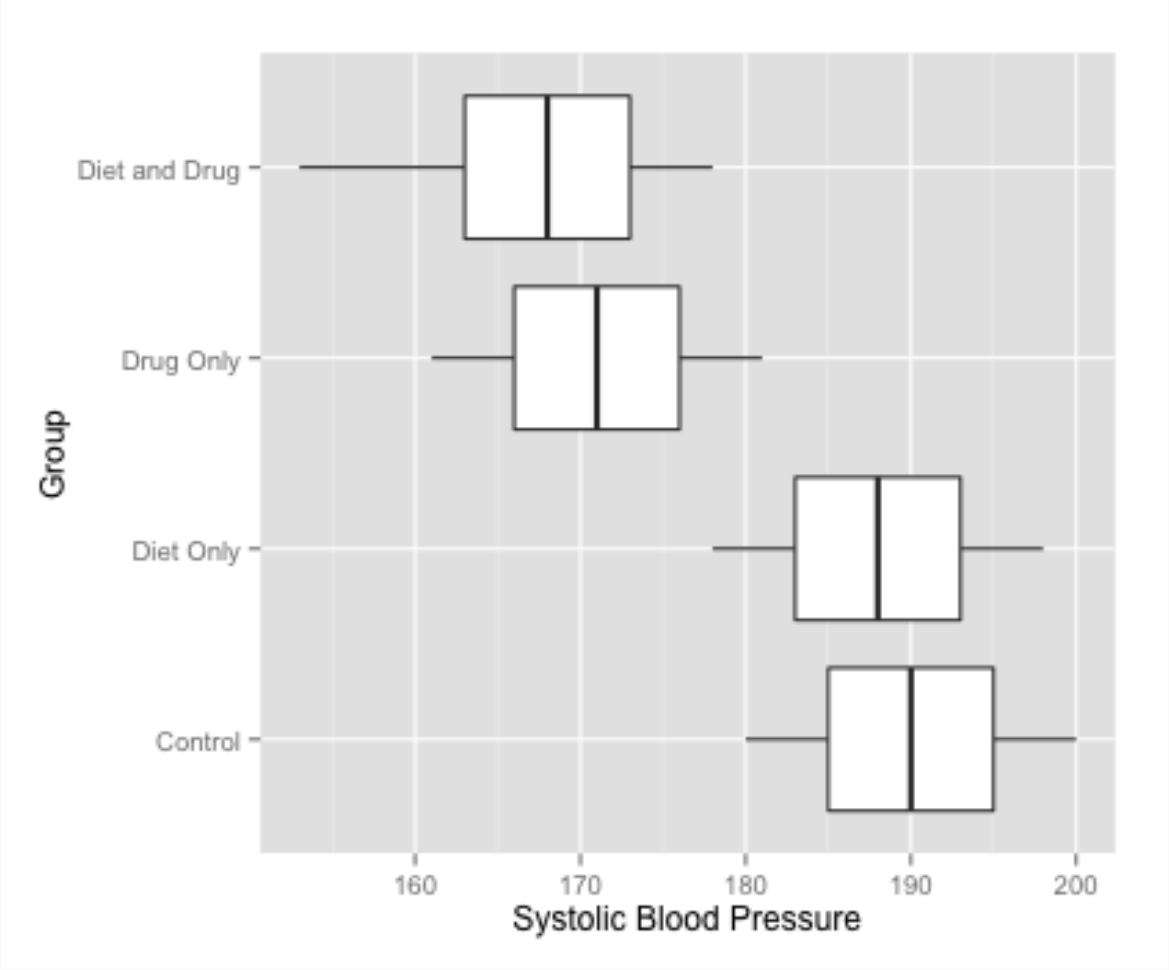
- Note how $\bar{y}$ drops as Temperature increases
- Note how $\bar{y}$ drops as Humidity rises
- Note that at 70^0 there is a difference of 0 between Low/High humidity
- Note that at 80^0 there is a difference of 10 between Low/High humidity
- Note that at 90^0 there is a difference of 20 between Low/High humidity
- Note that at Low Humidity there is a difference of 0 for every 10^0 rise in temperature
- Note that at High Humidity there is a difference of 10 for every 10^0 rise in temperature
- Now the effects of Factor A (temperature) differ for levels of Factor B (humidity)
- Interaction: Effect of one Factor depends upon levels of the other Factor


Interaction
Humidity
100
90
80
70
60
50
70
75
80
85
90
Temperature

Another Example: Drugs, Diet, and Blood Pressure

Let us say we randomly assign each of 20 subjects to one of four conditions and then measure their systolic blood pressure. The recorded measurements are shown below:

Group			
Control	Diet	Drug	Diet & Drug
185	188	171	153
190	183	176	163
195	198	181	173
200	178	166	178
180	193	161	168
$\bar{y} = 190$	$\bar{y} = 188$	$\bar{y} = 171$	$\bar{y} = 167$


Diet and Drug
Drug Only
Diet Only
Control
Group
160
170
180
190
200
Systolic Blood Pressure

Hypotheses for Two-Way ANOVA

1. There is no main effect of Diet. That is,

   H_0: $\mu_{noDiet} = \mu_{yesDiet}$

   H_1: $\mu_{noDiet} \neq \mu_{yesDiet}$

2. There is no main effect of Drug. That is,

   H_0: $\mu_{noDrug} = \mu_{yesDrug}$

   H_1: $\mu_{noDrug} \neq \mu_{yesDrug}$

3. There is no interaction effect between Diet and Drug. That is,

   H_0: $\mu_{noDrug.noDiet} = \mu_{noDrug.yesDiet} = \mu_{yesDrug.noDiet} = \mu_{yesDrug.yesDiet}$

   H_1: $\mu_{noDrug.noDiet} \neq \mu_{noDrug.yesDiet} \neq \mu_{yesDrug.noDiet} \neq \mu_{yesDrug.yesDiet}$

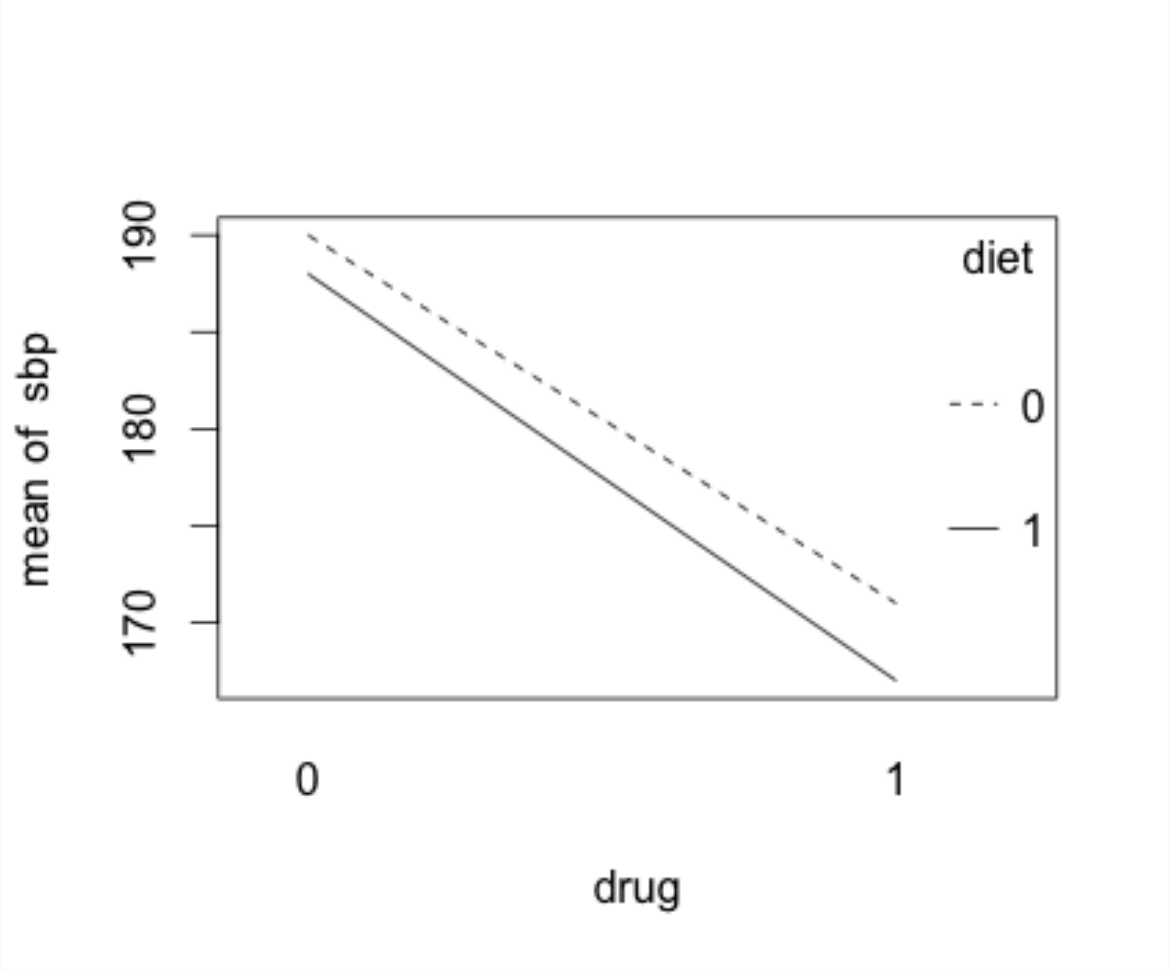
The Systolic Blood Pressure Example

```
> lm.0 <- lm(sbp ~ drug + diet + drug:diet, data=df)
> anova(lm.0)
Analysis of Variance Table

Response: sbp
          Df Sum Sq Mean Sq F value Pr(>F)
drug       1   2000    2000 28.5714 6.569e-05 ***
diet       1     45      45  0.6429   0.4344
drug:diet  1      5       5  0.0714   0.7927
Residuals 16   1120      70
---
Signif. codes: 0   ***   0.001   **   0.01   *   0.05   .   0.1       1
```

There is a statistically significant main effect of drug but not of diet

There is no statistically significant interaction effect of drug and diet


mean of sbp
190
180
170
0
1
drug
diet
0
1

The Lecture & Method Example

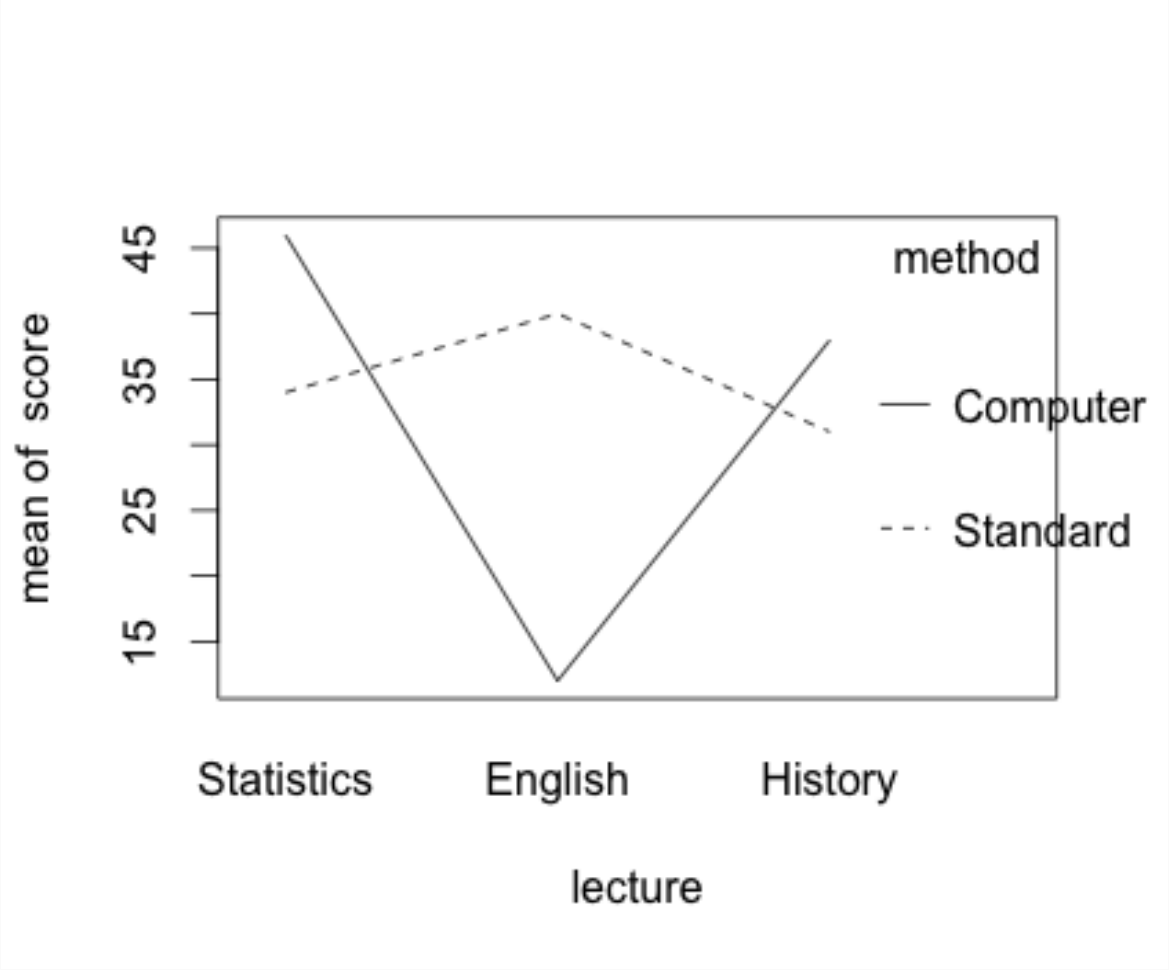
```
> lm.1 <- lm(score ~ lecture + method + lecture:method, data = df)
> anova(lm.1)
Analysis of Variance Table

Response: score
              Df Sum Sq Mean Sq F value Pr(>F)
lecture        2   1194   597.0 10.7374 0.000304 ***
method         1     81    81.0  1.4568 0.236866
lecture:method 2   2850  1425.0 25.6295 3.227e-07 ***
Residuals     30   1668    55.6
---
Signif. codes: 0   ***    0.001    **    0.01    *    0.05    .    0.1        1
```

There is a statistically significant main effect of lecture

There is no statistically significant main effect of method

There is a statistically significant interaction effect of lecture and method


mean of score
15
25
35
45
Statistics
English
History
lecture
method
Computer
Standard

Repeated Measures ANOVA

Repeated Measures Designs

1. In a test of over the counter headache medicines, each participant is given four medicines – Tylenol, Advil, Generic A, and Gingko – and medicinal effects measured four times per participant
2. Another design, called *profile analysis* involves comparing scores on different tests, with each participant completing each different test
   For example, men and women complete three scales, each designed to measure a different aspect of cultural sensitivity
   Note here you have two sources of potential "effects" – the different scales and the respondent's gender
3. A third design involves measuring some aspect of each participant at different times
   For example, with statewide standardized testing it is now possible to have every student in Ohio take the Ohio Achievement Assessments in Reading and in Mathematics in grades three through eight

- The test statistic is still the $F-ratio$
- There is one difference however – here the availability of multiple measures per subject better allows us to eliminate variance caused by subject-to-subject differences
- In particular, the F-ratio now becomes:

$$\frac{\text{Between-Groups variance - individual differences}}{\text{Within-groups variance - individual differences}}$$

- In the numerator, the fact that the same subjects are being measured more than once means subject-to-subject differences are automatically being eliminated
- In the denominator we will decompose Within-groups variance into that due to chance alone, versus that due to subject-to-subject differences
- As such the ANOVA model is fit in two stages
  1. Split total variance into Between-groups versus Within-groups variances
  2. Split Within-groups variances into those due to individual differences versus those due to sheer chance

Table 1: Effect of drug on amount of time (in seconds) a stimulus is endured

Person	Placebo	Drug A	Drug B	Drug C	Person Totals	
A	3	4	6	7	P = 20	n = 5
B	0	3	3	6	P = 12	k = 4
C	2	1	4	5	P = 12	N = 20
D	0	1	3	4	P = 8	G = 60
E	0	1	4	3	P = 8	$\sum y^2 = 262$
	T = 5 SS = 8	T = 10 SS = 8	T = 20 SS = 6	T = 25 SS = 10		

Note: T is the total for each group; SS = sum of squares; G = grand total

$$SS_{total} = \sum y^2 - \frac{G^2}{N} = 262 - \frac{3600}{20} = 262 - 180 = 82$$

$$SS_{Between-groups} = \sum \left(\frac{T^2}{n} \right) - \frac{G^2}{N} = 5 + 20 + 80 + 125 - 180 = 50$$

$$SS_{Within-groups} = 8 + 8 + 6 + 10 = 32$$

For the second stage we decompose the usual denominator

$$SS_{Between-persons} = \sum\left(\frac{P^2}{n}\right) - \frac{G^2}{N} = 100+36+36+16+16-180 = 24$$

$$SS_{error} = SS_{Within-groups} - SS_{Between-persons} = 32-24 = 8$$

The df Total = $N-1 = 20-1 = 19$; the df for Between-persons is $n-1 = 5-1 = 4$

The df Between-groups $= k-1 = 4-1 = 3$; the df Within-groups $= N-k = 20-4 = 16$

The df error $= (N-k)-(n-1) = 16-4 = 12$

So $MS_{Between-groups} = \frac{SS_{Between-groups}}{df_{Between-groups}} = \frac{50}{3} = 16.67$ and $MS_{error} = \frac{SS_{error}}{df_{error}} = \frac{8}{12} = 0.67$

$$F = \frac{MS_{Between-groups}}{MS_{error}} = \frac{16.67}{0.67} = 24.88$$

The p-value is < 0.01 so we reject H_0; data suggest mean endurance times are not equal across the four groups

```
> aov.1 = aov(response ~ group + Error(person/group), data = repeated)
> summary(aov.1)

Error: person
          Df Sum Sq Mean Sq F value Pr(>F)
Residuals 4     24       6

Error: person:group
          Df Sum Sq Mean Sq F value Pr(>F)
group      3     50  16.667     25 1.9e-05 ***
Residuals 12      8   0.667
---
Signif. codes: 0   ***   0.001   **   0.01   *   0.05   .   0.1       1
```

Advantages/Disadvantages of Repeated Measures Designs

- Advantages:
  1. Each participant serves as his/her own control
  2. Fewer participants are needed to achieve the same power as other designs because person-to-person differences are eliminated
  3. Most powerful design for analyzing change over time (e.g., value-added models)
- Disadvantages:
  1. Practice effects (subjects learn and this learning confounds measurement)
  2. Differential carry-over effects (some residual effects from the previous condition spills over into the next measurement)

Fixed versus Random Effects

Fixed Effects

Fixed effects refer to instances where the results are only generalizable to treatments that were a part of the research design; any treatment that was not explicitly tested is not covered by the ANOVA results

Example

Suppose you are a biostatistician at a small clinic with a total of five doctors. Clinic administrators are interested in answering the question: Does patient satisfaction vary significantly from one doctor to another? A number of patients of each doctor are asked to complete a survey, and the survey results are used in a statistical analysis that aims to answer the question.

This scenario calls for a one-way fixed-effects ANOVA because the administrators are interested in drawing conclusions about these five doctors. Since these five doctors are the only doctors, drawing conclusions about these doctors means drawing conclusions about the clinic as a whole. But this would not be true if the clinic had many doctors.

Random Effects

Random effects refer to situations where he treatments were randomly sampled from the population and hence the ANOVA results can be generalized to the population.

Example

Suppose you are a biostatistician at a large clinic with many doctors. Clinic administrators are interested in answering the question: Does patient satisfaction vary significantly from one doctor to another? Five doctors are selected at random, and a number of patients of each doctor are asked to complete a survey.

This scenario calls for a one-way random-effects ANOVA because the administrators are interested in drawing conclusions not about the five doctors selected but about the clinic as a whole. That is, the five doctors in the study are part of a larger population, and the population, rather than these five doctors or any five doctors, are of interest. The results from a fixed-effects analysis would not generalize to the clinic as a whole.

In random-effects ANOVA the groups enrolled in the study are not of primary interest since the goal is generalizing to the larger population

As such, random-effects ANOVA focuses on breaking down the total variance into two components

1. Within groups variance (σ^2)
2. Between groups variance (σ_A^2)

Example

The walking stick Timema cristinae is a wingless herbivorous insect that lives on plants in chaparral habitats in California. In a study of the insect's adaptations to different plant species, researchers measured a variety of traits using digital photographs of specimens collected from a study site. They used a computer to measure various traits in the photographs. Because they were concerned about measurement error they took two separate photographs of each specimen after some gap in time. Very often the second set of measurements were different from the first set, indicating measurement error. How large was the measurement error compared with the real variation among individuals in the trait?

Repeatability: $\frac{s_A^2}{s_A^2 + MS_{error}}$, where where $s_A^2 = \frac{MS_{(Between-groups)} - MS_{error}}{n}$

and $n =$ no. of measurements in each group

Repeatability measures the similarity of repeated measures on the same subject.

Repeatability = 0 implies a large role of measurement error and repeatability = 1 implies no measurement error

```
> walkingstickAnova <- lme(fixed = femurLength ~ 1, random = ~ 1|specimen, data = walkingstick)

> anova(walkingstickAnova)
            numDF denDF F-value p-value
(Intercept)     1    25 1021.889 <.0001

> walkingstickVarcomp <- VarCorr(walkingstickAnova)
> walkingstickVarcomp
specimen = pdLogChol(1)
            Variance    StdDev
(Intercept) 0.0010539182 0.03246411
Residual    0.0003559996 0.01886795

> varAmong <- as.numeric( walkingstickVarcomp[1,1] )
> varWithin <- as.numeric( walkingstickVarcomp[2,1] )
> repeatability <- varAmong / (varAmong + varWithin)
> repeatability
[1] 0.7475033

> 0.0010539182/(0.0010539182 + 0.0003559996)
[1] 0.7475033
```

Some Cautions

- If the ANOVA is only being carried out to estimate repeatability, then normality becomes less of an issue but the assumption of homogeneity of variances does still have to be met
- Note that the repeatability measure is sensitive to the nature of the sample used to estimate it. For example, if the sample is homogeneous (that is the between subject variance is very small), then the within subject variance will be proportionally larger and repeatability will be low. In other words it's all relative.
- Random-effects ANOVA models are fit using maximum likelihood techniques which, as a rule, need large samples to be really accurate
- By assumption groups were randomly sampled from the population