

Model Comparison and Inference, Hierarchical Models

February 14, 2022

Regression vs ANOVA

- Regression: find coefficients for $y = \beta_0 + \beta_1 x_1 + \dots$ that minimize the sum of squared differences between predicted y (\hat{y}) and actual y
- ANOVA: determine whether the estimate of variability in a sample based on group membership differs from the overall variability
- When x_i are discrete these two problems are mathematically equivalent
- Regression can also handle x_i that are continuous variables

Additive model approach to ANOVA

A	B	C	D
62	63	68	56
60	67	66	62
63	71	71	60
59	64	67	61
	65	68	63
	66	68	64
			63
			59

$\bar{y}_A=61$ $\bar{y}_B=66$ $\bar{y}_C=68$ $\bar{y}_D=61$
 $\bar{\bar{y}}=64$

24 df

=

A	B	C	D
64	64	64	64
64	64	64	64
64	64	64	64
64	64	64	64
	64	64	64
	64	64	64
			64
			64

grand mean

1 df

+

A	B	C	D
-3	+2	+4	-3
-3	+2	+4	-3
-3	+2	+4	-3
-3	+2	+4	-3
	+2	+4	-3
	+2	+4	-3
			-3
			-3

column effect

228

3 df

+

A	B	C	D
+1	-3	0	-5
-1	+1	-2	+1
+2	+5	+3	-1
-2	-2	-1	0
	-1	0	+2
	0	0	+3
			+2
			-2

error

112

20 df

$$F_{3,20} = \frac{76}{5.6} = 13.57$$



$$\sigma_{\tau}^2 = \frac{228}{3} = 76$$

$$\sigma_{\varepsilon}^2 = \frac{112}{20} = 5.6$$

```

> lm(obj_time ~ group, data = data1) %>% anova()
Analysis of Variance Table

Response: obj_time
          Df Sum Sq Mean Sq F value    Pr(>F)
group      3    228    76.0   13.571 4.658e-05 ***
Residuals 20    112     5.6
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> |

```

Regression with categorical <- dummy coding

```
> data1 %>% select(group, obj_time, group_B:group_D) %>% print(n=Inf)
# A tibble: 24 x 5
  group obj_time group_B group_C group_D
  <chr>   <dbl>   <dbl>   <dbl>   <dbl>
1 A         62      0      0      0
2 A         60      0      0      0
3 A         63      0      0      0
4 A         59      0      0      0
5 B         63      1      0      0
6 B         67      1      0      0
7 B         71      1      0      0
8 B         64      1      0      0
9 B         65      1      0      0
10 B        66      1      0      0
11 C         68      0      1      0
12 C         66      0      1      0
13 C         71      0      1      0
14 C         67      0      1      0
15 C         68      0      1      0
16 C         68      0      1      0
17 D         56      0      0      1
18 D         62      0      0      1
19 D         60      0      0      1
20 D         61      0      0      1
21 D         63      0      0      1
22 D         64      0      0      1
23 D         63      0      0      1
24 D         59      0      0      1
```

Regression with categorical <- dummy coding

```
> lm(obj_time ~ group_B + group_C + group_D, data = data1) %>% summary()

Call:
lm(formula = obj_time ~ group_B + group_C + group_D, data = data1)

Residuals:
    Min       1Q   Median       3Q      Max
-5.00  -1.25   0.00   1.25   5.00

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.100e+01  1.183e+00  51.554 < 2e-16 ***
group_B      5.000e+00  1.528e+00   3.273 0.003803 **
group_C      7.000e+00  1.528e+00   4.583 0.000181 ***
group_D      2.991e-15  1.449e+00   0.000 1.000000
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.366 on 20 degrees of freedom
Multiple R-squared:  0.6706,    Adjusted R-squared:  0.6212
F-statistic: 13.57 on 3 and 20 DF,  p-value: 4.658e-05
```

```
> lm(obj_time ~ group, data = data1) %>% summary()

Call:
lm(formula = obj_time ~ group, data = data1)

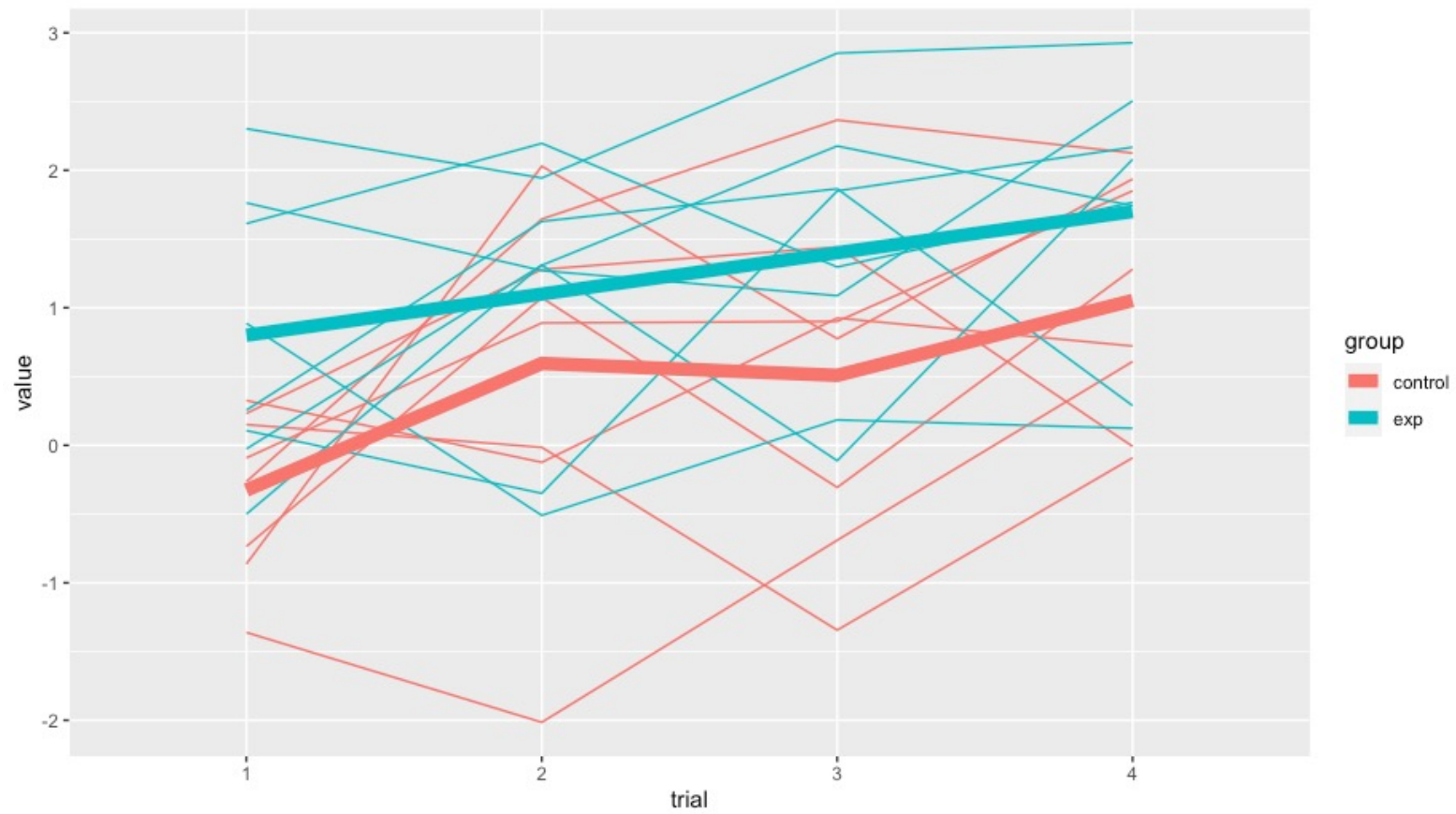
Residuals:
    Min       1Q   Median       3Q      Max
-5.00  -1.25   0.00   1.25   5.00

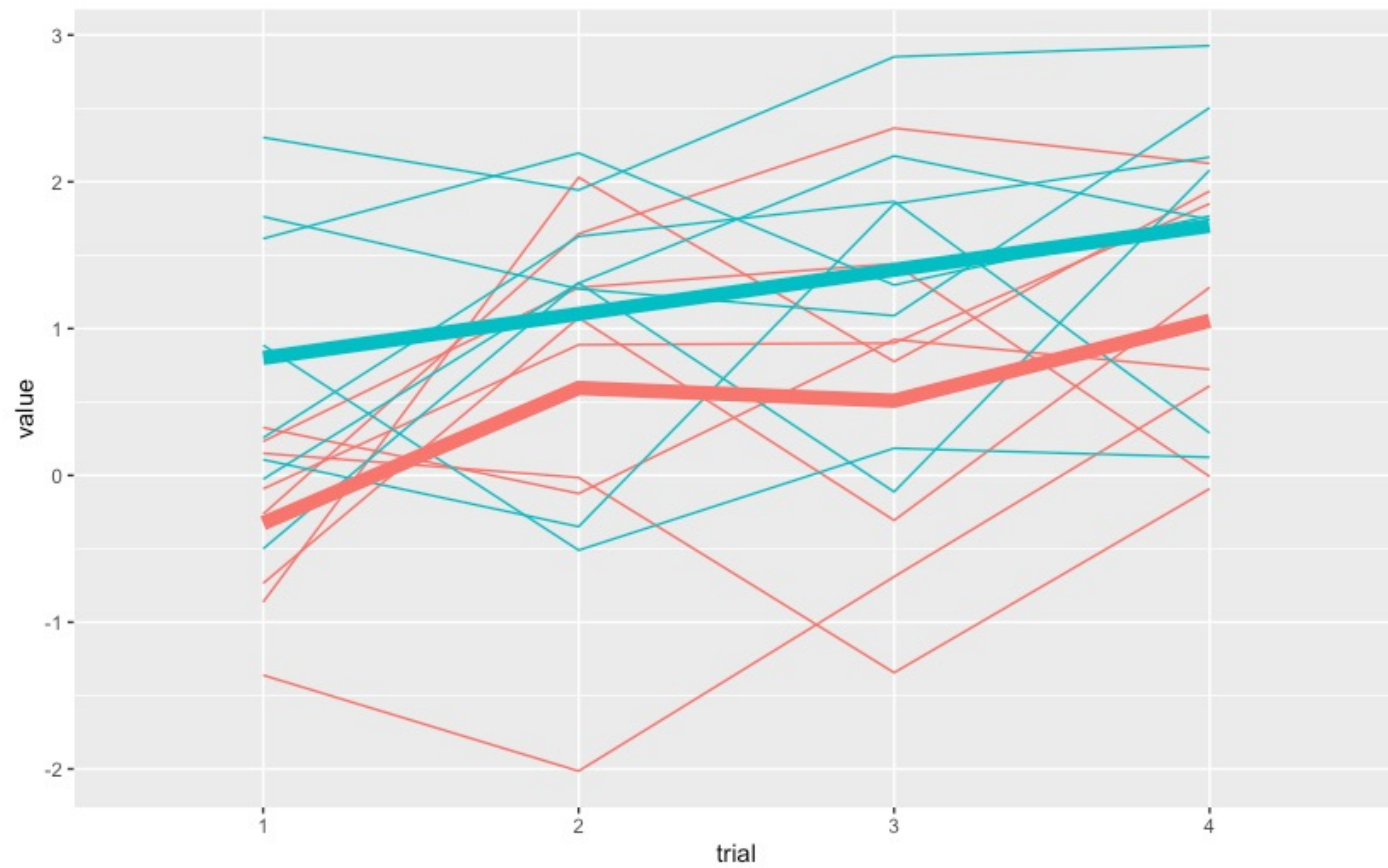
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.100e+01  1.183e+00  51.554 < 2e-16 ***
groupB      5.000e+00  1.528e+00   3.273 0.003803 **
groupC      7.000e+00  1.528e+00   4.583 0.000181 ***
groupD      2.991e-15  1.449e+00   0.000 1.000000
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.366 on 20 degrees of freedom
Multiple R-squared:  0.6706,    Adjusted R-squared:  0.6212
F-statistic: 13.57 on 3 and 20 DF,  p-value: 4.658e-05
```

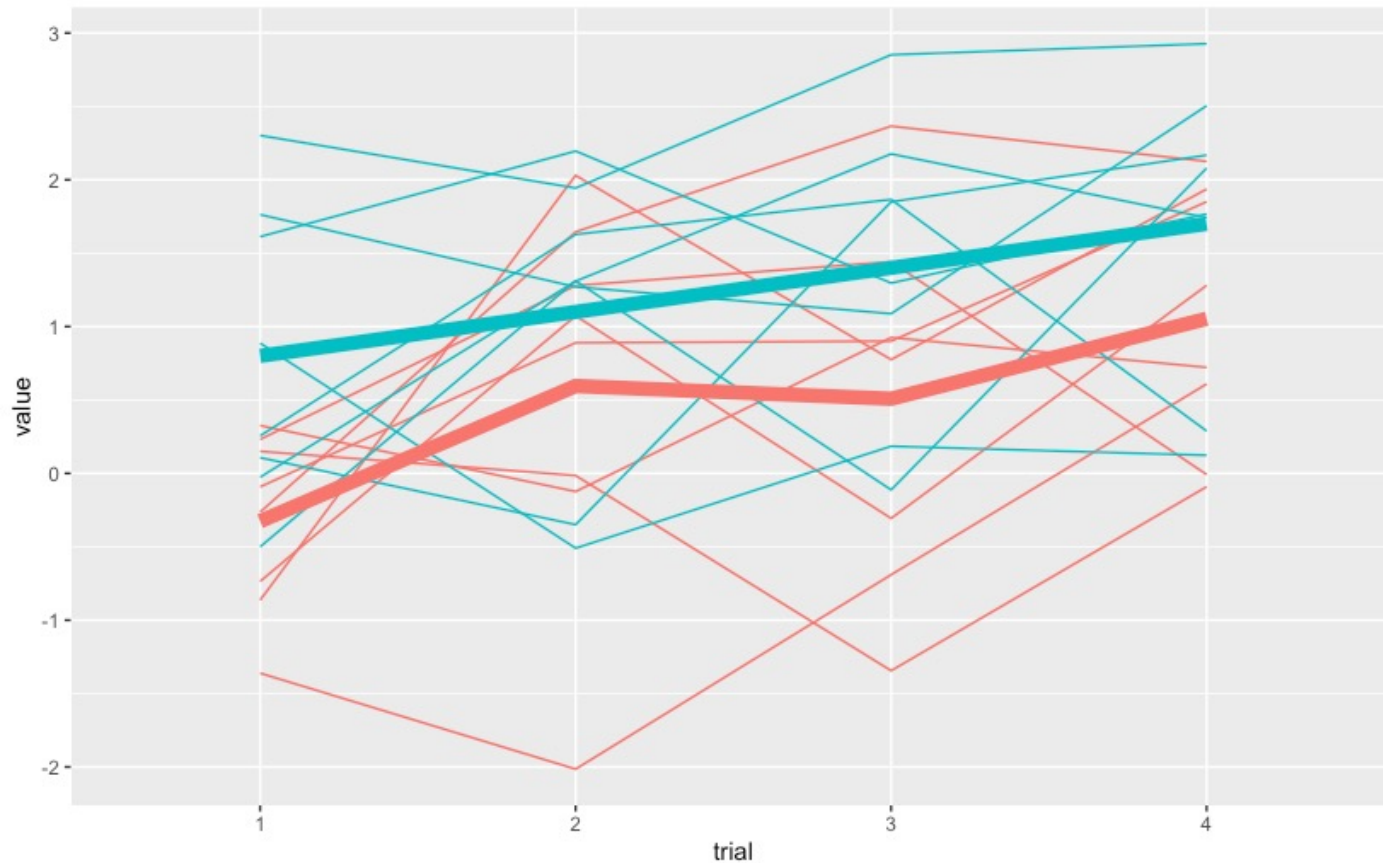
Repeated measures

- Conceptually this is just another kind of dummy coding
- Critical to model the dependence in the residuals correctly – expect observations from same subject / unit of analysis to be more similar to each other than observations from different subjects





main effect of group:
collapse across trial



main effect of group:
collapse across trial

main effect of trial:
difference across trials irrespective
of group membership

	linear trend	quadratic trend	cubic trend
trial == 1	-3	-1	-1
trial == 2	-1	1	3
trial == 3	1	1	-3
trial == 4	3	-1	1

dummy coding for polynomial contrasts

cannot get any of these from linear combination of others (add/subtract)

dot products of these vectors = 0 – they are uncorrelated (orthogonal)

$$-3 * -1 + -1 * 1 + 1 * 1 + 3 * -1 = 3 + -1 + 1 + -3 = 0$$

elements of each add to zero = net influence on predicted value is 0

```
> lmer(value ~ group + trial + (1|subject_ID), data = long_data) %>% summary()
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: value ~ group + trial + (1 | subject_ID)
Data: long_data

REML criterion at convergence: 173.1

Scaled residuals:
    Min       1Q   Median       3Q      Max
-1.92537 -0.61932  0.07798  0.76722  1.52659

Random effects:
 Groups      Name      Variance Std.Dev.
subject_ID (Intercept) 0.3770   0.6140
Residual          0.6372   0.7982
Number of obs: 64, groups: subject_ID, 16

Fixed effects:
              Estimate Std. Error    df t value Pr(>|t|)
(Intercept)  0.45774    0.25892 14.00000   1.768 0.098860 .
groupexp     0.79226    0.36617 14.00000   2.164 0.048272 *
trial.L      0.78827    0.19956 45.00000   3.950 0.000272 ***
trial.Q     -0.09424    0.19956 45.00000  -0.472 0.639053
trial.C      0.18340    0.19956 45.00000   0.919 0.362982
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
      (Intr) gropxp tril.L tril.Q
groupexp -0.707
trial.L   0.000 0.000
trial.Q   0.000 0.000 0.000
trial.C   0.000 0.000 0.000 0.000
```

```
> lm(value ~ group + linear_trend + quad_trend + cubic_trend, data = long_data) %>% summary()

Call:
lm(formula = value ~ group + linear_trend + quad_trend + cubic_trend,
    data = long_data)

Residuals:
    Min       1Q   Median       3Q      Max
-2.4662 -0.5974  0.1425  0.7078  1.8072

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.45774    0.17634   2.596  0.01189 *
groupexp     0.79226    0.24938   3.177  0.00237 **
linear_trend  0.17626    0.05576   3.161  0.00248 **
quad_trend   0.04712    0.12469   0.378  0.70688
cubic_trend   0.04101    0.05576   0.735  0.46499
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9975 on 59 degrees of freedom
Multiple R-squared:  0.2604,    Adjusted R-squared:  0.2102
F-statistic: 5.192 on 4 and 59 DF,  p-value: 0.001188
```

```
> lmer(value ~ group + trial + (1|subject_ID), data = long_data) %>% summary()
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: value ~ group + trial + (1 | subject_ID)
Data: long_data

REML criterion at convergence: 173.1

Scaled residuals:
    Min       1Q   Median       3Q      Max
-1.92537 -0.61932  0.07798  0.76722  1.52659

Random effects:
 Groups      Name      Variance Std.Dev.
subject_ID (Intercept) 0.3770   0.6140
Residual          0.6372   0.7982
Number of obs: 64, groups: subject_ID, 16

Fixed effects:
              Estimate Std. Error    df t value Pr(>|t|)
(Intercept)  0.45774    0.25892 14.00000   1.768 0.098860 .
groupexp     0.79226    0.36617 14.00000   2.164 0.048272 *
trial.L      0.78827    0.19956 45.00000   3.950 0.000272 ***
trial.Q     -0.09424    0.19956 45.00000  -0.472 0.639053
trial.C      0.18340    0.19956 45.00000   0.919 0.362982
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
      (Intr) gropxp tril.L tril.Q
groupexp -0.707
trial.L   0.000 0.000
trial.Q   0.000 0.000 0.000
trial.C   0.000 0.000 0.000 0.000
```

```
> lm(value ~ group + linear_trend + quad_trend + cubic_trend, data = long_data) %>% summary()

Call:
lm(formula = value ~ group + linear_trend + quad_trend + cubic_trend,
    data = long_data)

Residuals:
    Min       1Q   Median       3Q      Max
-2.4662 -0.5974  0.1425  0.7078  1.8072

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.45774    0.17634   2.596  0.01189 *
groupexp     0.79226    0.24938   3.177  0.00237 **
linear_trend  0.17626    0.05576   3.161  0.00248 **
quad_trend   0.04712    0.12469   0.378  0.70688
cubic_trend   0.04101    0.05576   0.735  0.46499
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9975 on 59 degrees of freedom
Multiple R-squared:  0.2604,    Adjusted R-squared:  0.2102
F-statistic: 5.192 on 4 and 59 DF,  p-value: 0.001188
```

lmer vs lm: estimate of group effect is the same but p-values differ

Repeated measures

```
> lm(value ~ group + linear_trend + quad_trend + cubic_trend, data = long_data_3) %>% summary()

Call:
lm(formula = value ~ group + linear_trend + quad_trend + cubic_trend,
    data = long_data_3)

Residuals:
    Min       1Q   Median       3Q      Max
-2.2397 -0.7877  0.2430  0.7054  1.8047

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.42479    0.18326   2.318  0.0241 *
groupexp      1.61337    0.25924   6.224 6.17e-08 ***
linear_trend  0.31601    0.05796   5.452 1.11e-06 ***
quad_trend    0.01156    0.12962   0.089  0.9292
cubic_trend   0.03494    0.05794   0.603  0.5489
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.02 on 57 degrees of freedom
(2 observations deleted due to missingness)
Multiple R-squared:  0.541,    Adjusted R-squared:  0.5087
F-statistic: 16.79 on 4 and 57 DF,  p-value: 3.786e-09
```

```
> lmer(value ~ group * trial + (1|subject_ID), data = long_data_3) %>% summary()
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: value ~ group * trial + (1 | subject_ID)
Data: long_data_3

REML criterion at convergence: 164.4

Scaled residuals:
    Min       1Q   Median       3Q      Max
-2.02285 -0.66654  0.08559  0.65051  1.43750

Random effects:
 Groups      Name                Variance Std.Dev.
subject_ID (Intercept) 0.3768    0.6138
Residual              0.6331    0.7957
Number of obs: 62, groups:  subject_ID, 16

Fixed effects:
            Estimate Std. Error    df t value Pr(>|t|)
(Intercept)    0.4366    0.2603 14.2954   1.678 0.115162
groupexp        1.6172    0.3681 14.2954   4.394 0.000583 ***
trial.L         0.9246    0.2825 40.3892   3.273 0.002187 **
trial.Q        -0.1462    0.2873 40.6252  -0.509 0.613691
trial.C         0.3101    0.2921 40.8460   1.062 0.294667
groupexp:trial.L  1.0087    0.4064 40.6252   2.482 0.017282 *
groupexp:trial.Q  0.2539    0.4064 40.6252   0.625 0.535648
groupexp:trial.C -0.2619    0.4064 40.6252  -0.645 0.522869
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```


Repeated measures

```
> lm(value ~ group + linear_trend + quad_trend + cubic_trend, data = long_data_3) %>% summary()

Call:
lm(formula = value ~ group + linear_trend + quad_trend + cubic_trend,
    data = long_data_3)

Residuals:
    Min       1Q   Median       3Q      Max
-2.2397 -0.7877  0.2430  0.7054  1.8047

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.42479    0.18326   2.318  0.0241 *
groupexp      1.61337    0.25924   6.224 6.17e-08 ***
linear_trend  0.31601    0.05796   5.452 1.11e-06 ***
quad_trend   0.01156    0.12962   0.089  0.9292
cubic_trend   0.03494    0.05794   0.603  0.5489
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.02 on 57 degrees of freedom
(2 observations deleted due to missingness)
Multiple R-squared:  0.541,    Adjusted R-squared:  0.5087
F-statistic: 16.79 on 4 and 57 DF,  p-value: 3.786e-09
```

```
> lmer(value ~ group * trial + (1|subject_ID), data = long_data_3) %>% summary()
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: value ~ group * trial + (1 | subject_ID)
Data: long_data_3

REML criterion at convergence: 164.4

Scaled residuals:
    Min       1Q   Median       3Q      Max
-2.02285 -0.66654  0.08559  0.65051  1.43750

Random effects:
 Groups      Name      Variance Std.Dev.
subject_ID (Intercept) 0.3768   0.6138
Residual              0.6331   0.7957
Number of obs: 62, groups: subject_ID, 16

Fixed effects:
            Estimate Std. Error    df t value Pr(>|t|)
(Intercept)    0.4366    0.2603 14.2954   1.678 0.115162
groupexp        1.6172    0.3681 14.2954   4.394 0.000583 ***
trial.L         0.9246    0.2825 40.3892   3.273 0.002187 **
trial.Q        -0.1462    0.2873 40.6252  -0.509 0.613691
trial.C         0.3101    0.2921 40.8460   1.062 0.294667
groupexp:trial.L  1.0087    0.4064 40.6252   2.482 0.017282 *
groupexp:trial.Q  0.2539    0.4064 40.6252   0.625 0.535648
groupexp:trial.C -0.2619    0.4064 40.6252  -0.645 0.522869
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Missing data: lm throws out the whole case
lmer adjusts for different N at each trial

Getting those p-values

- `emmeans` package
 - <https://timmastny.rbind.io/blog/tests-pairwise-categorical-mean-emmeans-contrast/>
 - <https://aosmith.rbind.io/2019/03/25/getting-started-with-emmeans/>
 - <https://cran.r-project.org/web/packages/emmeans/vignettes/basics.html>
- Generates estimated marginal means for effects – weighted means based on cell sizes with appropriate confidence intervals

emmeans

```
> ld3model <- lmer(value ~ group * trial + (1|subject_ID), data = long_data_3)
>
> anova(ld3model)
Type III Analysis of Variance Table with Satterthwaite's method
```

	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)	
group	15.1837	15.1837	1	13.986	25.5964	0.0001748	***
trial	31.3660	10.4553	3	40.340	17.6254	1.821e-07	***
group:trial	5.1617	1.7206	3	40.340	2.9005	0.0465440	*

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

emmeans

```
> ld3model <- lmer(value ~ group * trial + (1|subject_ID), data = long_data_3)
>
> anova(ld3model)
Type III Analysis of Variance Table with Satterthwaite's method
```

	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)	
group	15.1837	15.1837	1	13.986	25.5964	0.0001748	***
trial	31.3660	10.4553	3	40.340	17.6254	1.821e-07	***
group:trial	5.1617	1.7206	3	40.340	2.9005	0.0465440	*

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

emmeans

```
> ld3model <- lmer(value ~ group * trial + (1|subject_ID), data = long_data_3)
>
> anova(ld3model)
Type III Analysis of Variance Table with Satterthwaite's method
      Sum Sq Mean Sq NumDF DenDF F value    Pr(>F)
group    15.1837  15.1837     1  13.986 25.5964 0.0001748 ***
trial    31.3660  10.4553     3  40.340 17.6254 1.821e-07 ***
group:trial  5.1617   1.7206     3  40.340  2.9005 0.0465440 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

if you are interested in where significant changes occur as trial factor increases, this is one way to get that information

```
> ld3model %>% emmeans(consec ~ trial | group, adjust = "none")
```

\$emmeans

group = control:

trial	emmean	SE	df	lower.CL	upper.CL
1	-0.542	0.337	40.0	-1.223	0.14
2	0.556	0.356	43.4	-0.162	1.27
3	0.395	0.337	40.0	-0.287	1.08
4	0.834	0.337	40.0	0.152	1.52

group = exp:

trial	emmean	SE	df	lower.CL	upper.CL
1	0.800	0.337	40.0	0.118	1.48
2	1.600	0.337	40.0	0.918	2.28
3	2.400	0.337	40.0	1.718	3.08
4	3.391	0.356	43.4	2.673	4.11

Degrees-of-freedom method: kenward-roger

Confidence level used: 0.95

\$contrasts

group = control:

contrast	estimate	SE	df	t.ratio	p.value
2 - 1	1.098	0.402	40.6	2.732	0.0093
3 - 2	-0.162	0.402	40.6	-0.402	0.6898
4 - 3	0.439	0.385	40.0	1.140	0.2612

group = exp:

contrast	estimate	SE	df	t.ratio	p.value
2 - 1	0.800	0.385	40.0	2.077	0.0442
3 - 2	0.800	0.385	40.0	2.077	0.0442
4 - 3	0.991	0.402	40.6	2.466	0.0180

Degrees-of-freedom method: kenward-roger

Hierarchical comparisons

- `anova(model1, model2)` will evaluate the difference in residual sums of squares between two models
- a significant test statistic indicates that the second model accounts for more variance / "fits better" than the first model
- **this can be useful in cases where you want to evaluate complex interactions or whether some factors have an influence "over and above" that of other factors**
- this is also a better approach when you are analyzing GLMs

Hierarchical comparisons

```
> ld3model <- lmer(value ~ group * trial + (1|subject_ID), data = long_data_3)
> ld3model_reduced <- lmer(value ~ group + trial + (1|subject_ID), data = long_data_3) # main effects, no interactions
>
> anova(ld3model, ld3model_reduced)
refitting model(s) with ML (instead of REML)
Data: long_data_3
Models:
ld3model_reduced: value ~ group + trial + (1 | subject_ID)
ld3model: value ~ group * trial + (1 | subject_ID)

```

	npars	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
ld3model_reduced	7	176.08	190.97	-81.038	162.08			
ld3model	10	172.92	194.19	-76.460	152.92	9.1557	3	0.02729 *

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Hierarchical comparisons

```
> ld3model <- lmer(value ~ group * trial + (1|subject_ID), data = long_data_3)
> ld3model_reduced <- lmer(value ~ group + trial + (1|subject_ID), data = long_data_3) # main effects, no interactions
>
> anova(ld3model, ld3model_reduced)
refitting model(s) with ML (instead of REML)
Data: long_data_3
Models:
ld3model_reduced: value ~ group + trial + (1 | subject_ID)
ld3model: value ~ group * trial + (1 | subject_ID)

```

	npars	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
ld3model_reduced	7	176.08	190.97	-81.038	162.08			
ld3model	10	172.92	194.19	-76.460	152.92	9.1557	3	0.02729 *

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Akaike information criterion

Bayesian information criterion

Log likelihood

Hierarchical comparisons

```
> ld3model <- lmer(value ~ group * trial + (1|subject_ID), data = long_data_3)
> ld3model_reduced <- lmer(value ~ group + trial + (1|subject_ID), data = long_data_3) # main effects, no interactions
>
> anova(ld3model, ld3model_reduced)
refitting model(s) with ML (instead of REML)
Data: long_data_3
Models:
ld3model_reduced: value ~ group + trial + (1 | subject_ID)
ld3model: value ~ group * trial + (1 | subject_ID)

```

	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
ld3model_reduced	7	176.08	190.97	-81.038	162.08			
ld3model	10	172.92	194.19	-76.460	152.92	9.1557	3	0.02729 *

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

number of parameters

Akaike information criterion

Bayesian information criterion

Log likelihood

difference in deviance ($-2 * \log \text{likelihood}$)
distributed as chi-squared