The General Linear Model – ANOVA Part 1 (Video 4)

Dr Andrew J. Stewart

E: drandrewjstewart@gmail.com
T: @ajstewart_lang
G: ajstewartlang

Factorial ANOVA

So far we have looked at ANOVA for designs when we have one factor which is between participants (i.e., each participant appears in only one condition), and for designs when we have one factor that is repeated measures (each participant appears in all conditions). These are examples of 1-way ANOVA.

Now we're going to look at factorial ANOVA - this is for cases where we have more than one factor and we might be interested in how the factors interact with each other. If we have two factors, we will build a 2-way ANOVA, three factors, a 3-way ANOVA etc.

Factorial ANOVA

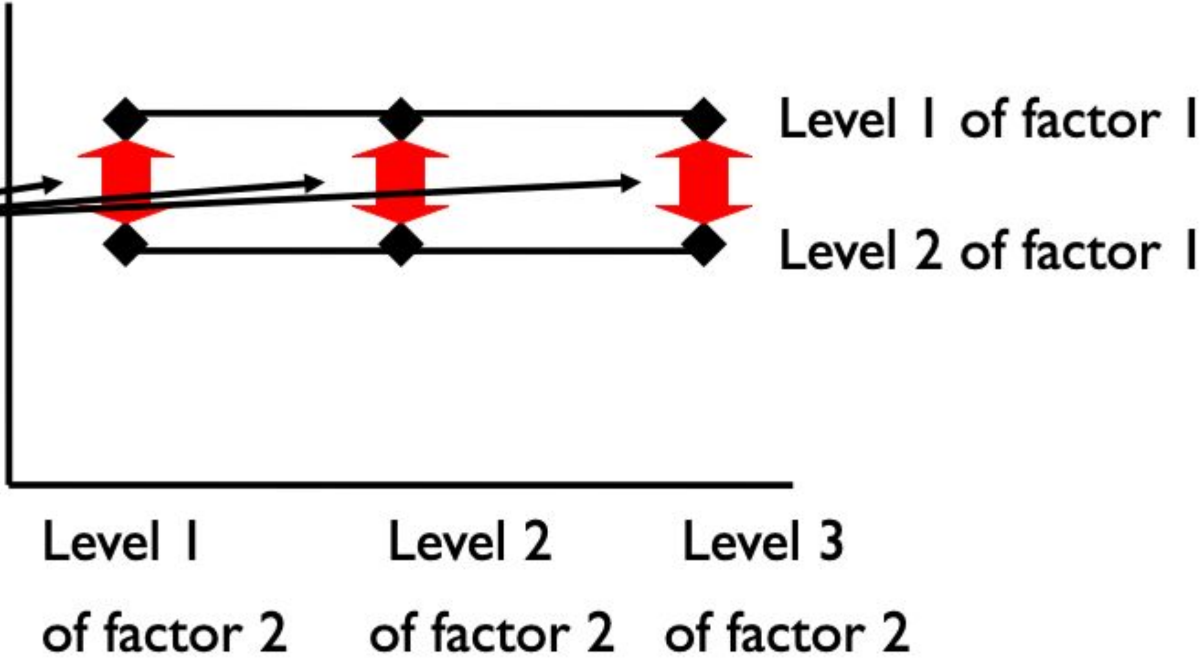
Imagine we have 2 factors. Factor 1 with two levels, Factor 2 with three. Our analysis might reveal a main effect of Factor 1 (i.e., a difference between the two levels), a main effect of Factor 2 (i.e., a difference between the three levels) or an interaction between the two…..

To examine what might be going on, we build a 2 x 3 ANOVA. The first number corresponds to the first factor and the number of levels it has (2) while the second corresponds to the second factor and the number of levels it has (3).

Following are examples of ‘perfect’ patterns of effects...

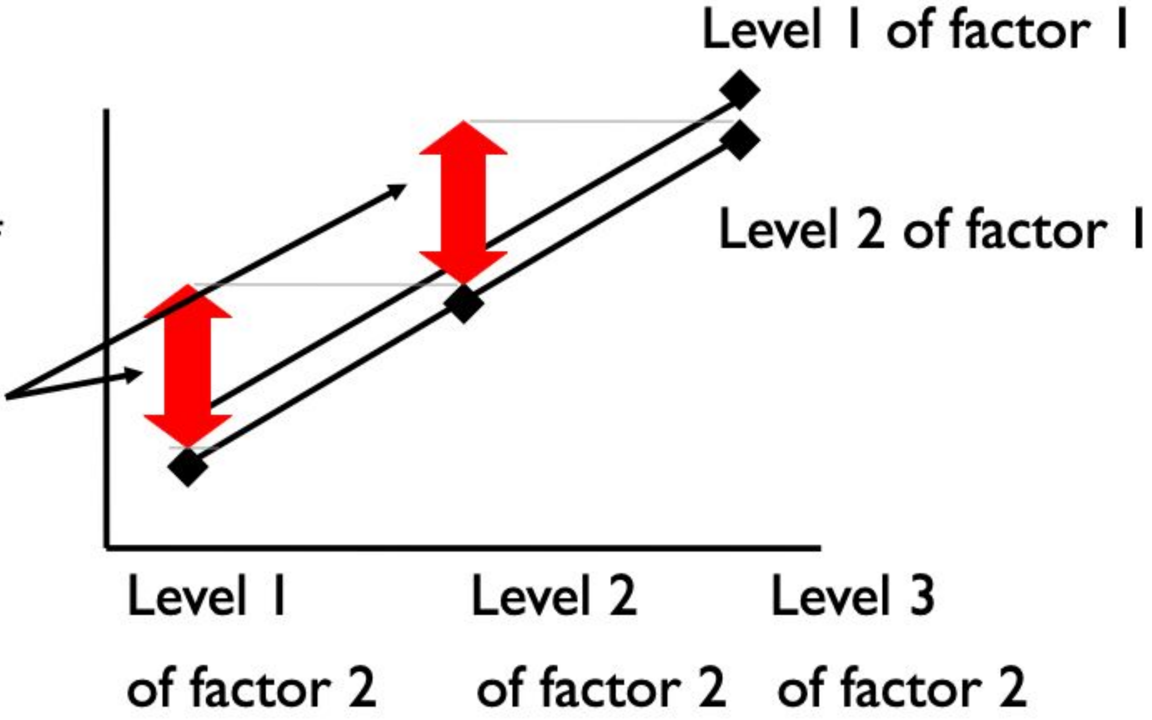
Main effect of Factor 1, no main effect of Factor 2 and no interaction

The differences between levels I and 2 of Factor I are all significant and are of the same magnitude.



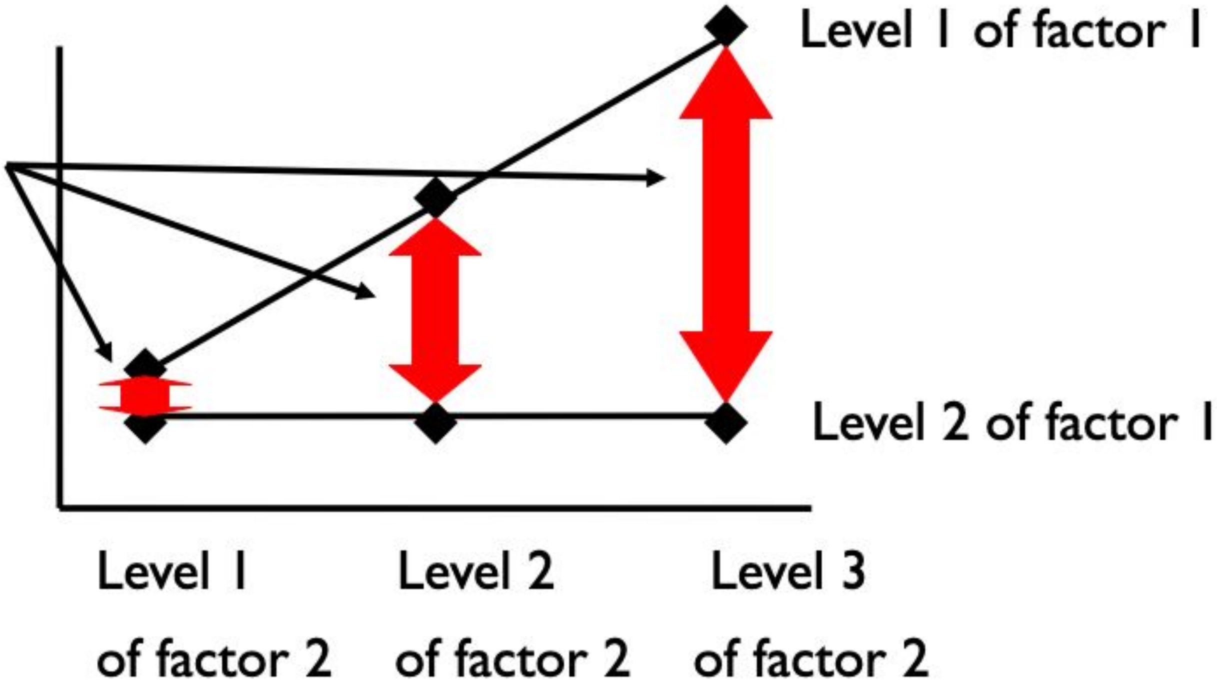
No main effect of Factor 1, main effect of Factor 2 and no interaction

The differences between levels 1 & 2 and 2 & 3 of Factor 2 are all significant and are of the same magnitude. There are no significant differences between levels 1 and 2 of Factor 1.



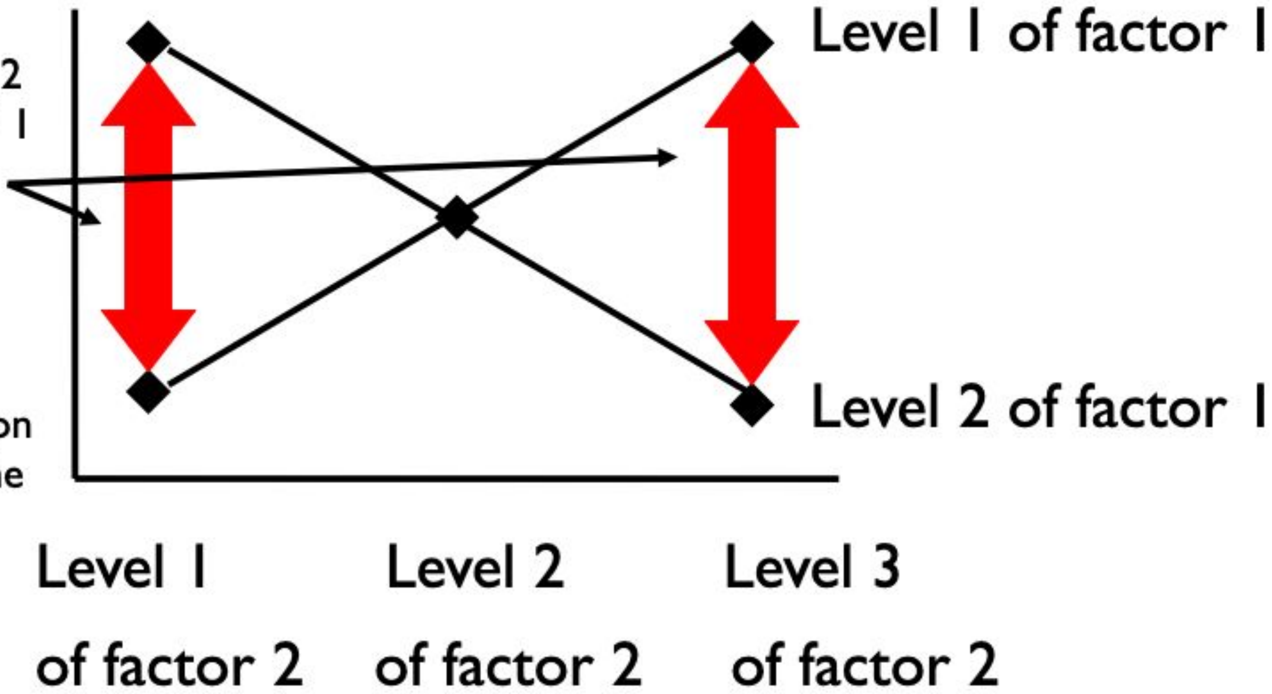
Main effect of Factor 1, main effect of Factor 2 and an interaction

The differences between the two levels of factor 1 change as a function of factor 2.



No main effect of Factor 1, no main effect of Factor 2 but an interaction

The difference between levels 1 & 2 of Factor 1 at Level 1 of Factor 2 is different from the same difference at Levels 2 and 3 of Factor 2. This is a crossover interaction as the polarity of the difference flips.



Example Factorial ANOVA

Imagine the case where we're interested in the effect of positive vs. negative contexts on how quickly (in milliseconds) people respond to positive vs negative sentences. We think there might be a priming effect (i.e., people are quicker to respond to positive sentences after positive contexts vs. after negative contexts - and vice versa).

So, we have two factors, each with two levels. This is what's known as a full factorial design where every subject participates in every condition.

Reading in our Data

```
factorial_data <-
read_csv("https://raw.githubusercontent.com/ajstewartlang/11_glm_anova_pt1/
master/data/factorial_data.csv")
head(factorial_data)

# A tibble: 6 x 5
  Subject  Item    RT Sentence Context
    <dbl> <dbl> <dbl> <chr>    <chr>
1       1     3  1270 Positive Negative
2       1     7   739 Positive Negative
3       1    11   982 Positive Negative
4       1    15  1291 Positive Negative
5       1    19  1734 Positive Negative
6       1    23  1757 Positive Negative
```

Tidying our Data

```
factorial_data_tidied <- factorial_data %>%
  mutate(Sentence = factor(Sentence), Context = factor(Context))
head(factorial_data_tidied)

# A tibble: 6 x 5
  Subject  Item    RT Sentence Context
    <dbl> <dbl> <dbl> <fct>    <fct>
1       1     3  1270 Positive Negative
2       1     7   739 Positive Negative
3       1    11   982 Positive Negative
4       1    15  1291 Positive Negative
5       1    19  1734 Positive Negative
6       1    23  1757 Positive Negative
```

Summarising our Data

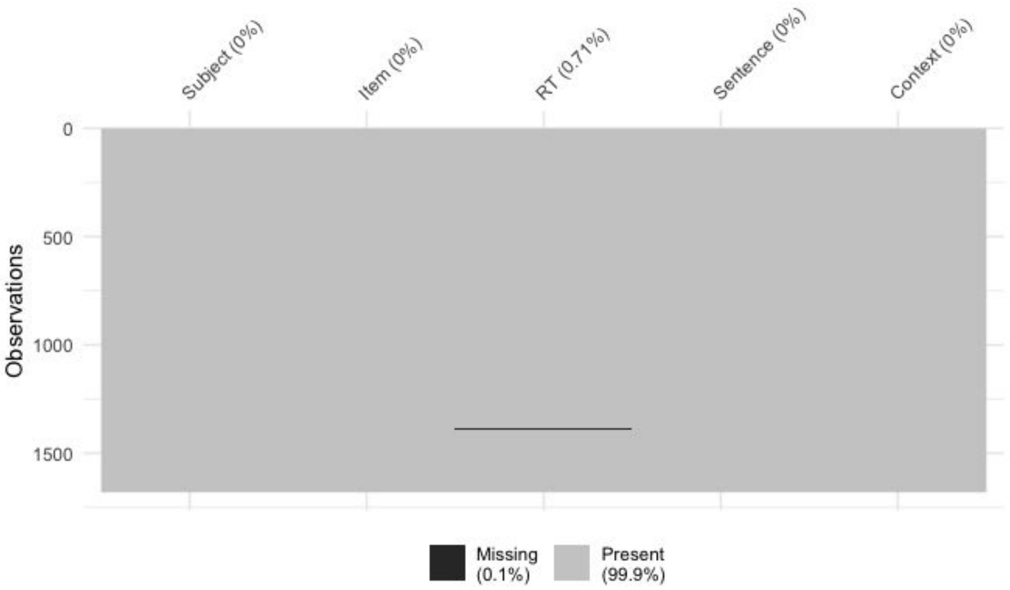
```
factorial_data_tidied %>%
  group_by(Context, Sentence) %>%
  summarise(mean_rt = mean(RT), sd_rt = sd(RT))

# A tibble: 4 x 4
# Groups:   Context [2]
  Context  Sentence mean_rt sd_rt
  <fct>    <fct>      <dbl> <dbl>
1 Negative Negative   1474.  729.
2 Negative Positive     NA    NA
3 Positive Negative     NA    NA
4 Positive Positive   1579.  841.
```

What’s happening here?

Do we have missing data?

```
vis_miss(factorial_data_tidied)
```

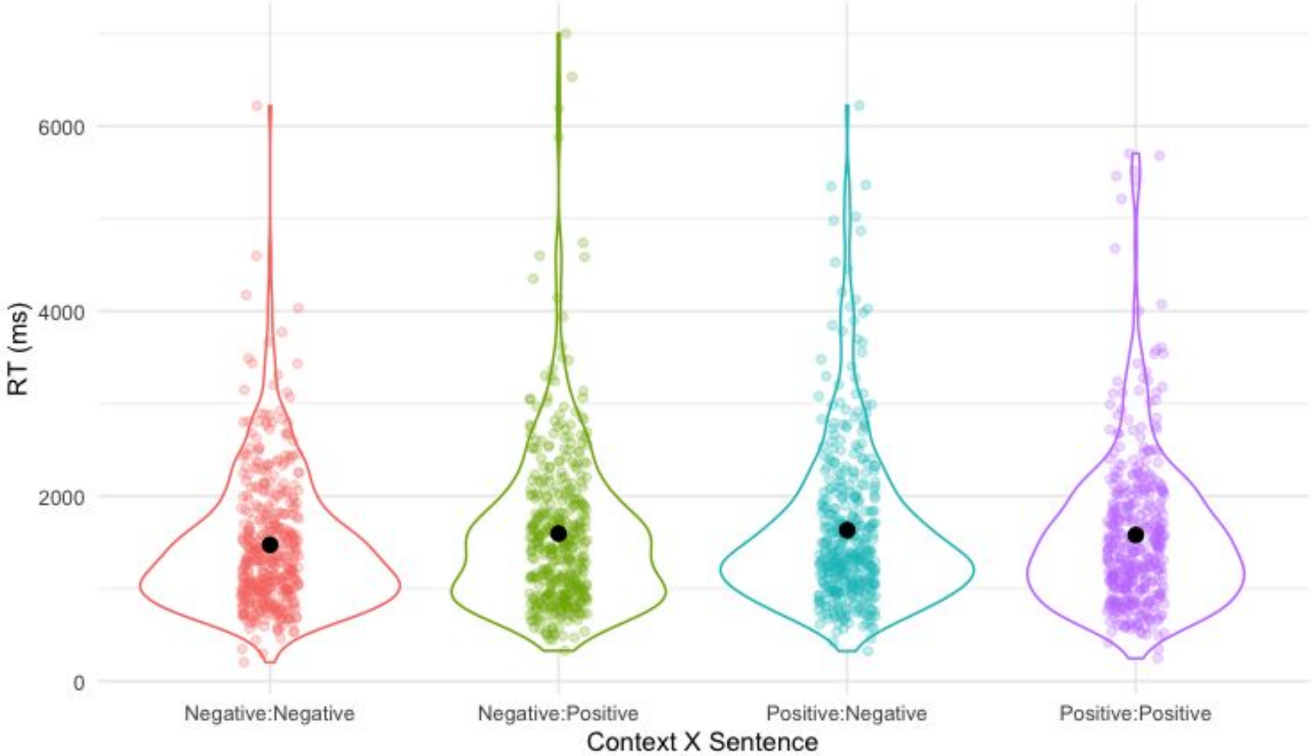


Ignoring Missing Data

```
factorial_data_tidied %>%
  group_by(Context, Sentence) %>%
  summarise(mean_rt = mean(RT, na.rm = TRUE), sd_rt = sd(RT, na.rm =
TRUE))

# A tibble: 4 x 4
# Groups:   Context [2]
  Context  Sentence mean_rt sd_rt
  <fct>    <fct>      <dbl> <dbl>
1 Negative Negative   1474.  729.
2 Negative Positive   1595.  887.
3 Positive Negative   1633.  877.
4 Positive Positive   1579.  841.
```

Visualising our Data



Modelling our Data – F1

```
model_subjects <- aov_4(RT ~ Context * Sentence + (1 + Context * Sentence |
Subject), data = factorial_data_tidied, na.rm = TRUE)
anova(model_subjects)

Anova Table (Type 3 tests)

Response: RT
                 num Df den Df    MSE      F       ges  Pr(>F)
Context               1     59  90195 3.1767 0.0060231 0.07984 .
Sentence              1     59 124547 0.6283 0.0016524 0.43114
Context:Sentence      1     59  93889 4.5967 0.0090449 0.03616 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Modelling our Data – F2

```
model_items <- aov_4(RT ~ Context * Sentence + (1 + Context * Sentence |
Item), data = factorial_data_tidied, na.rm = TRUE)
anova(model_items)

Anova Table (Type 3 tests)

Response: RT
                 num Df den Df    MSE      F       ges   Pr(>F)
Context               1     27  39844 4.0013 0.0080150 0.05561 .
Sentence              1     27 203164 0.1221 0.0012553 0.72951
Context:Sentence      1     27  40168 5.7687 0.0116070 0.02346 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Interpreting our Model

```
emmeans(model_subjects, pairwise ~ Context * Sentence, adjust = "none")
$emmeans
 Context  Sentence emmean   SE  df lower.CL upper.CL
 Negative Negative   1474 57.8 138     1360     1588
 Positive Negative   1628 57.8 138     1514     1742
 Negative Positive   1595 57.8 138     1481     1709
 Positive Positive   1579 57.8 138     1465     1693

Warning: EMMs are biased unless design is perfectly balanced
Confidence level used: 0.95

$contrasts
 contrast                              estimate   SE  df t.ratio p.value
 Negative Negative - Positive Negative   -153.9 55.4 118 -2.779  0.0064
 Negative Negative - Negative Positive   -120.9 60.3 116 -2.004  0.0474
 Negative Negative - Positive Positive   -105.2 59.8 115 -1.759  0.0813
 Positive Negative - Negative Positive     33.0 59.8 115  0.551  0.5824
 Positive Negative - Positive Positive     48.7 60.3 116  0.807  0.4213
 Negative Positive - Positive Positive     15.7 55.4 118  0.284  0.7772
```

These are the two key comparisons.

Interpreting our Model

We conducted a 2 (Context: Positive vs. Negative) x 2 (Sentence: Positive vs. Negative) repeated measures ANOVA to investigate the influence of Context valence on reaction times to Sentences of Positive or Negative valence. The ANOVA revealed no effect of Sentence ($F < 1$), no effect of Context ($F(1, 59) = 3.18$, $p = .080$, $\eta G^2 = .006$), but an interaction between Sentence and Context ($F(1, 59) = 4.60$, $p = .036$, $\eta G^2 = .009$).

The interaction was interpreted by conducting Bonferroni-corrected pairwise companions. These comparisons revealed that the interaction was driven by Negative Sentences being processed faster in Negative vs. Positive Contexts (1,474 ms. vs. 1,628 ms., $t(118) = 2.78$, $p = .012$) while Positive Sentences were read at similar speeds in Negative vs. Positive Contexts (1,595 ms. vs. 1,579 ms., $t(118) = .284$, $p = 1$).

Reporting ANOVA

- Say what type of ANOVA it was, say what factors you had (and with labels for each level).
- Report the results of main effects first, then interactions.
- Report F values, exact *p*-values, effect sizes, and confidence intervals.
- Remember to interpret interactions further - such as with contrasts or pairwise comparisons.
- When you have main effects, say which direction the effect goes.
- Avoid sillies - e.g., mixing up < and > or saying $p = .000$