DATA2002

Two-factor ANOVA with interactions

Garth Tarr

THE UNIVERSITY OF SYDNEY

Two-factor ANOVA: more than one treatment

Interaction plots

Post hoc tests

Two-factor ANOVA

Multi-factor ANOVA

- In some experiments there is more than one *factor* that might affect the response.
- We can think of such an experiment as a "big one-way ANOVA" where each *combination of factors* is a different "treatment".
- Interest lies in determining not only if there are differences between each treatment, but other questions, such as whether
  - each factor has an effect;
  - the effect of one factor is the same across all levels of the other factor(s).
- We shall focus on the case of **two factors**.

General setup

We have observations

$$\{y_{ijk}\colon\ i = 1, \ldots, a;\ j = 1, \ldots, b;\ k = 1, \ldots, n\}$$

where y_{ijk} is the k-th observation receiving the treatment combination corresponding to level i of factor A and level j of factor B.

- There are n observations receiving each treatment combination.
- The **full model** is that y_{ijk} is the value taken by $Y_{ijk} \sim N(\mu_{ij}, \sigma^2)$.

Examples

Poisons and antidotes

- In a famous paper Box and Cox (1964) analyse an experiment where one of each of **3 poisons** and **4 antidotes** were administered to a sample of 4 animals, giving 48 observations all together (4 observations on each of the 12 treatment combinations).
    - The response was **survival time**.
    - They showed the **reciprocal** of the survival time was an appropriate transformation to use on the response.
- The aim of the study was to determine how each `antidote` affected survival in the presence of each `poison`.
- This is a "3-by-4 factorial experiment".

Poisons and antidotes

The data can be found in the **BHH2** package (Barrios, 2016).

```
# install.packages("BHH2")
data("poison.data", package = "BHH2")
glimpse(poison.data)
```

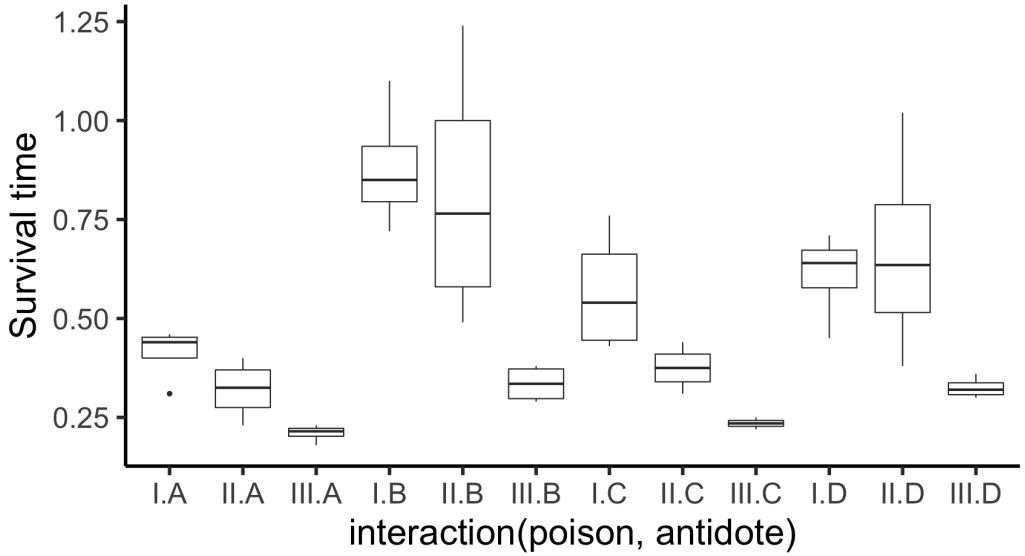
```
## Rows: 48
## Columns: 3
## $ poison <fct> I, I, I, I, II, II, II, II, III, III, III, …
## $ treat  <fct> A, A, A, A, A, A, A, A, A, A, A, A, B, B, B…
## $ y      <dbl> 0.31, 0.45, 0.46, 0.43, 0.36, 0.29, 0.40, 0…
```

```
# rename treat as antidote to avoid confusion with the general term "treatment"
poison.data = poison.data %>%
  rename(antidote = treat)
glimpse(poison.data)
```

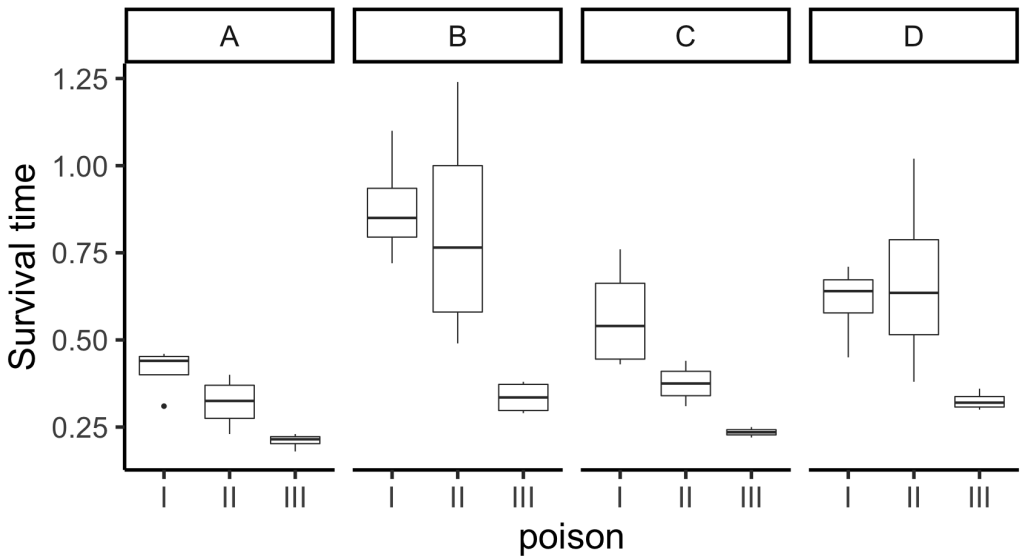
```
## Rows: 48
## Columns: 3
## $ poison   <fct> I, I, I, I, II, II, II, II, III, III, III…
## $ antidote <fct> A, A, A, A, A, A, A, A, A, A, A, A, B, B,…
## $ y        <dbl> 0.31, 0.45, 0.46, 0.43, 0.36, 0.29, 0.40,…
```

Visualising the data

```
poison.data %>% ggplot() +
  aes(x = interaction(poison, antidote), y = y)
  geom_boxplot() +
  theme_classic(base_size = 30) +
  labs(y = "Survival time")
```

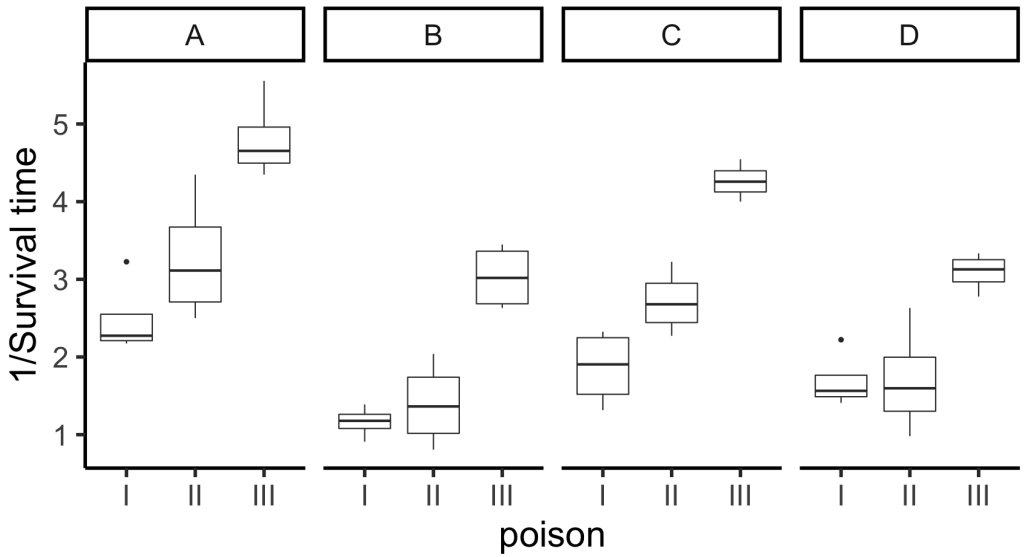



```
poison.data %>% ggplot() +
  aes(x = poison, y = y) +
  geom_boxplot() +
  theme_classic(base_size = 30) +
  facet_wrap(~antidote, ncol = 4) +
  labs(y = "Survival time")
```



The variance is much less for groups with lower means. A transformation is clearly needed here to satisfy normality, equal variance assumptions.

```
poison.data = poison.data %>%
  mutate(inv_survival = 1/y)
poison.data %>% ggplot() +
  aes(x = poison, y = inv_survival) +
  geom_boxplot() +
  theme_classic(base_size = 30) +
  facet_wrap(~antidote, ncol = 4) +
  labs(y = "1/Survival time")
```



- The reciprocal transformation is much better;
  - while spreads still differ somewhat, they don't get systematically smaller with smaller mean.

Paper planes

- Mackisack (1994) has the results from an experiment where statistics students launched paper planes in a controlled environment, controlling for various factors, including
  - `Paper` quality: 1 (80gsm) and 2 (50gsm);
  - `Plane` design: 1 (High-performance glider) and 2 (Incredibly simple glider).
- The response was distance travelled in mm.
- Four "flights" were conducted at each of the 4 treatment combinations.
- Do either `Paper` or `Plane` (or both) have any impact on the distance "flown"?
- This is a "2-by-2 factorial experiment".

Data source: http://www.statsci.org/data/oz/planes.html

```
planes = read_tsv("https://raw.githubusercontent.com/DATA2002/data/master/planes.txt")
glimpse(planes)
```

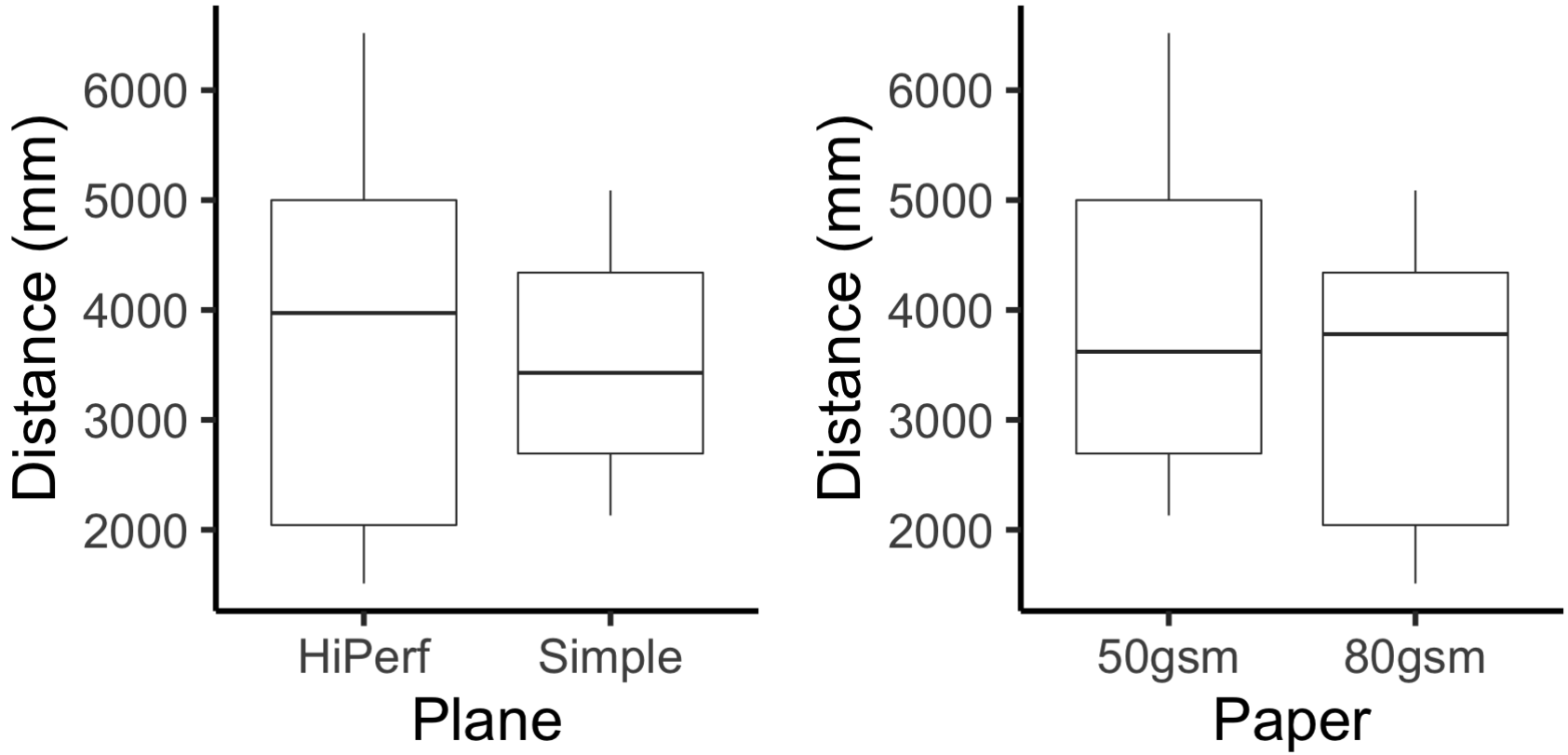
```
## Rows: 16
## Columns: 5
## $ Distance <dbl> 2160, 1511, 4596, 3706, 3854, 1690, 5088,…
## $ Paper    <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2,…
## $ Angle    <dbl> 1, 1, 1, 1, 2, 2, 2, 2, 1, 1, 1, 1, 2, 2,…
## $ Plane    <dbl> 1, 1, 2, 2, 1, 1, 2, 2, 1, 1, 2, 2, 1, 1,…
## $ Order    <dbl> 12, 11, 8, 6, 4, 2, 1, 7, 3, 9, 14, 5, 16…
```

```
planes = planes %>%
  mutate(
    Paper = case_when(
      Paper == 1 ~ "80gsm",
      Paper == 2 ~ "50gsm"
    ),
    Plane = case_when(
      Plane == 1 ~ "HiPerf",
      Plane == 2 ~ "Simple"
    )
  )
```

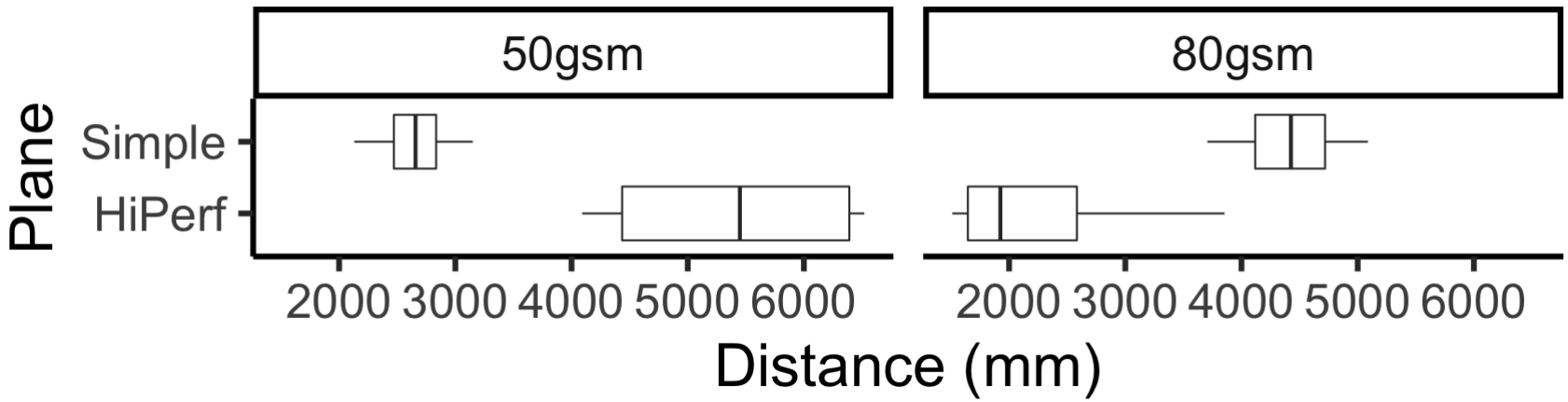
```
glimpse(planes)
```

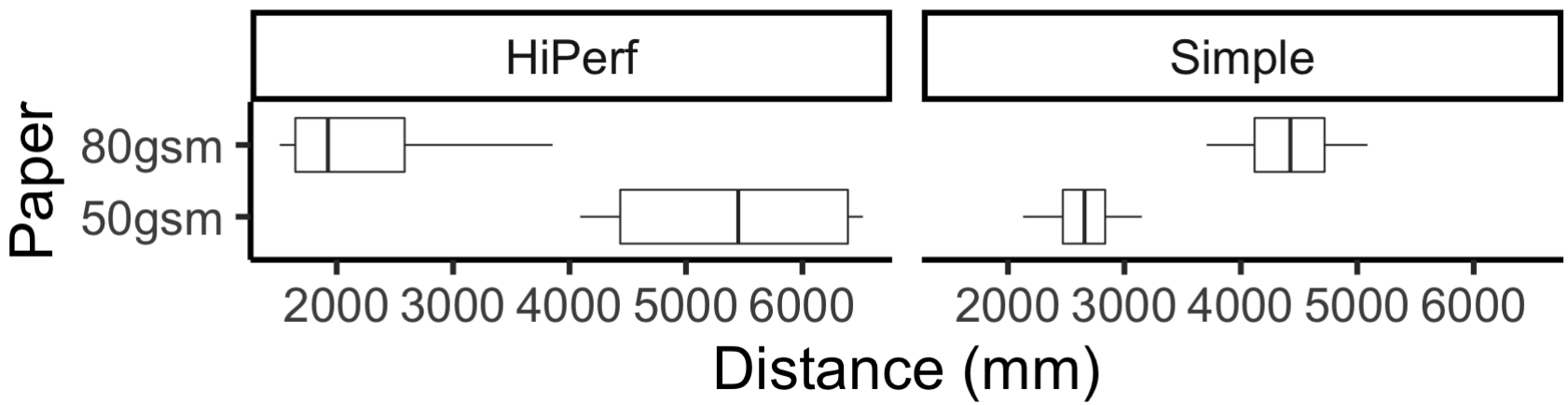
```
## Rows: 16
## Columns: 5
## $ Distance <dbl> 2160, 1511, 4596, 3706, 3854, 169
## $ Paper    <chr> "80gsm", "80gsm", "80gsm", "80gsm
## $ Angle    <dbl> 1, 1, 1, 1, 2, 2, 2, 2, 1, 1, 1,
## $ Plane    <chr> "HiPerf", "HiPerf", "Simple", "Si
## $ Order    <dbl> 12, 11, 8, 6, 4, 2, 1, 7, 3, 9,
```

```
p1 = ggplot(planes, aes(x = Plane, y = Distance)) +
  geom_boxplot() + labs(y = "Distance (mm)")
p2 = ggplot(planes, aes(x = Paper, y = Distance)) +
  geom_boxplot() + labs(y = "Distance (mm)")
gridExtra::grid.arrange(p1, p2, ncol = 2)
```



```
p3 = p1 + facet_wrap(~ Paper) + coord_flip()
p4 = p2 + facet_wrap(~ Plane) + coord_flip()
gridExtra::grid.arrange(p3, p4, ncol = 1)
```





General setup

- For each $i = 1, \ldots, a$, $j = 1, \ldots, b$, $k = 1, \ldots, n$, y_{ijk} is the k-th observation receiving level i of factor A and level j of factor B and is modelled as the value taken by the random variable,

$$Y_{ijk} \sim N(\mu_{ij}, \sigma^2) \,,$$

  and all random variables are assumed independent.

Various "questions"

- We might "ask various questions" about the structure of the μ_{ij}'s (i.e. *test various null hypotheses*), e.g.:
  - H_0: $\mu_{ij} \equiv \mu$ (i.e. straight 1-way ANOVA between all treatment combinations)
  - H_0: $\mu_{ij} \equiv \mu_i$ for all j (so factor B has no effect)
  - H_0: $\mu_{ij} \equiv \mu_j$ for all i (so factor A has no effect)
- We can also ask if an **interaction** is present
  - H_0: $\mu_{ij} = \mu + \alpha_i + \gamma_j$ for some constants μ, α_i ($i = 1, \ldots, a$), γ_j ($j = 1, \ldots, b$), so that the two factors combine **additively**.
  - If this is true then the adjustment for each level of factor A is *the same within each level of factor B*, and vice versa.
  - If this is **not** true, then we say there is some **interaction** between factors A and B.
    - This is (therefore) known as a "*test of no interaction*".

This is **not** a two-way ANOVA (in the sense of adjusting for blocks)

- This looks very similar to our earlier model where we were adjusting for blocks.
- However, it is **very different**.
- When adjusting for blocks:
  - the full model is $E(Y_{ij}) = \mu_{ij} = \mu + \alpha_i + \beta_j$ (i.e. treatments and blocks combine additively);
  - the null hypothesis is that $\mu_{ij} = \mu + \beta_j$ (i.e. all "treatment effects" α_i are zero);
  - the β_j's are **not** treatment effects, they are block effects.
- In **this** scenario,
  - the full model is $E(Y_{ijk}) = \mu_{ij}$ (i.e. unrestricted);
  - the null hypothesis is $\mu_{ij} = \mu + \alpha_i + \gamma_j$ (factor A and factor B effects combine additively);
  - both α_i and γ_j are related to treatment effects.

- We therefore use γ_j (i.e. a different Greek letter to β_j) to stress this difference.
- In many books
  - β_j is used for both: **bad!**;
  - this setup is called "two-way ANOVA with replicates": **also bad**;
    - better to call it "two-factor ANOVA".

Some theory!

Reparametrisation: contrasts!

- We introduce some new "parameters".
- Define
  - $\mu = \bar{\mu}_{\bullet\bullet}$ ("overall mean")
  - $\alpha_i = \bar{\mu}_{i\bullet} - \bar{\mu}_{\bullet\bullet}$ ("main effect due to i-th level of factor A")
  - $\gamma_j = \bar{\mu}_{\bullet j} - \bar{\mu}_{\bullet\bullet}$ ("main effect due to j-th level of factor B")
  - and the interaction effect between level i of factor A and level j of factor B:

$$(\alpha\gamma)_{ij} = \mu_{ij} - (\mu + \alpha_i + \gamma_j) = \mu_{ij} - \bar{\mu}_{i\bullet} - \bar{\mu}_{\bullet j} + \bar{\mu}_{\bullet\bullet}$$

- Then we may write

$$\mu_{ij} = \mu + \alpha_i + \gamma_j + (\alpha\gamma)_{ij} \; .$$

- The α_i's, γ_j's and $(\alpha\gamma)_{ij}$'s are all **contrasts** in the μ_{ij}'s!

Main effects: contrasts!

- Each α_i is a contrast in the μ_{ij}'s. For example,

$$\begin{aligned}\alpha_1 = \bar{\mu}_{1\bullet} - \bar{\mu}_{\bullet\bullet} &= \frac{1}{b}(\mu_{11} + \ldots + \mu_{1b}) - \frac{1}{ab}(\mu_{11} + \ldots + \mu_{ab}) \\ &= \left[\frac{1}{b} - \frac{1}{ab}\right] \underbrace{(\mu_{11} + \ldots + \mu_{1b})}_{b \text{ terms here}} - \frac{1}{ab}\underbrace{(\mu_{21} + \ldots + \mu_{ab})}_{(a-1)b \text{ terms here}}\end{aligned}$$

- This is of the form $\sum_{i=1}^{a}\sum_{j=1}^{b} c_{ij}\mu_{ij}$ where
  - b of the c_{ij}'s are equal to $\frac{1}{b} - \frac{1}{ab} = \frac{a-1}{ab}$;
  - the remaining $(a-1)b$ terms are equal to $-\frac{1}{ab}$.
  - these add to

$$b\frac{(a-1)}{ab} - (a-1)b\frac{1}{ab} = 0\,.$$

so α_1 is a contrast in the μ_{ij}'s!

This slide is only shown for the interests of mathematically inclined students.

Main effects: interpretation

- So each α_i in fact is a contrast in the (treatment combination) group means that measures in some overall sense how the means for level i of factor A differ from the overall average.
- In exactly the same way, each γ_j is a contrast (in the μ_{ij}'s) that compares (in some overall sense) level j of factor B to the overall average.

Interaction effects

- A similar (but more complicated) calculation can be used to show that each $(\alpha\gamma)_{ij}$ is also a contrast in the μ_{ij}'s.
- Each

$$(\alpha\gamma)_{ij} = \mu_{ij} - (\mu + \alpha_i + \gamma_j)$$

  compares a mean μ_{ij} to the corresponding "additive prediction" $\mu + \alpha_i + \gamma_j$.

  - If the factor levels *actually do* combine additively, then each such interaction (population) contrast is zero.
- Therefore a "test for no interaction" can be formulated as the following null hypothesis:

$$H_0\text{: } (\alpha\gamma)_{ij} = 0 \text{ for all } i, j.$$

Constraints: main effects degrees of freedom

- Note that (as with our earlier models) these "effects" satisfy certain constraints:
- Both sets of main effects add to zero:

$$\sum_{i=1}^{a} \alpha_i = \sum_{j=1}^{b} \gamma_j = 0\,.$$

- There are thus
  - $a-1$ "free" α_i's and
  - $b-1$ "free" γ_j's.
- That is to say, there are
  - $a-1$ degrees of freedom for the factor A main effects;
  - $b-1$ degrees of freedom for the factor B main effects.

The model specification, the constraints required to make the model "work" and the degrees of freedom are examinable!

Constraints: interaction effects degrees of freedom

- The interaction terms are such that for each fixed i,

$$\sum_{j=1}^{b}(\alpha\gamma)_{ij} = 0$$

  and vice versa for each fixed j.
- There are thus $(a-1)(b-1)$ "free" interaction effects (like in an a by b two-way contingency table where all row and column sums are fixed).
- That is to say, there are $(a-1)(b-1)$ degrees of freedom for the interaction effects.

Fitted values and residuals for each model

- We thus have two models:
  - the full model where μ_{ij} is "unrestricted";
  - the ("no interaction") null hypothesis where $\mu_{ij} = \mu + \alpha_i + \gamma_j$.
- Under each model, each observation y_{ijk} is decomposed into two parts:
  - a fitted value $\hat{\mu}_{ij}$;
  - a residual $y_{ijk} - \hat{\mu}_{ij}$.
- Under the full model $\hat{\mu}_{ij} = \bar{y}_{ij\bullet}$, the mean of the (i, j)-th **treatment combination**.
- Under the "no interaction" model, the fitted value is

$$\hat{\mu}_{ij} = \hat{\mu} + \hat{\alpha}_i + \hat{\gamma}_j = \bar{y}_{\bullet\bullet\bullet} + (\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}) + (\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet}) = \bar{y}_{i\bullet\bullet} + \bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet}$$

Residual sum of squares for each model

- So for the **full model**:
  - the (i,j,k)-th residual is $y_{ijk} - \bar{y}_{ij\bullet}$;
  - the **residual sum of squares** is

$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(y_{ijk} - \bar{y}_{ij\bullet}\right)^2.$$

- For the **no interaction model**:
  - the (i,j,k)-th residual is $y_{ijk} - \bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet j\bullet} + \bar{y}_{\bullet\bullet\bullet}$;
  - the **residual sum of squares** is

$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(y_{ijk} - \bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet j\bullet} + \bar{y}_{\bullet\bullet\bullet}\right)^2.$$

Decomposition

$$
\begin{aligned}
\text{Total SS} &= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk}-\bar{y}_{\bullet\bullet\bullet})^2 \\
&= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\Big[(\bar{y}_{i\bullet\bullet}-\bar{y}_{\bullet\bullet\bullet})+(\bar{y}_{\bullet j\bullet}-\bar{y}_{\bullet\bullet\bullet})+(y_{ijk}-\bar{y}_{i\bullet\bullet}-\bar{y}_{\bullet j\bullet}+\bar{y}_{\bullet\bullet\bullet})\Big]^2 \\
&= \underbrace{\sum_{i=1}^{a} bn(\bar{y}_{i\bullet\bullet}-\bar{y}_{\bullet\bullet\bullet})^2}_{\text{Factor A Sum Sq.}} + \underbrace{\sum_{j=1}^{b} an(\bar{y}_{\bullet j\bullet}-\bar{y}_{\bullet\bullet\bullet})^2}_{\text{Factor B Sum Sq.}} + \underbrace{\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk}-\bar{y}_{i\bullet\bullet}-\bar{y}_{\bullet j\bullet}+\bar{y}_{\bullet\bullet\bullet})^2}_{\text{no interaction model Res Sum Sq.}} + \text{ cross-product terms (all 0)} \\
&= \sum_{i=1}^{a} bn(\bar{y}_{i\bullet\bullet}-\bar{y}_{\bullet\bullet\bullet})^2 + \sum_{j=1}^{b} an(\bar{y}_{\bullet j\bullet}-\bar{y}_{\bullet\bullet\bullet})^2 + \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\Big[(\bar{y}_{ij\bullet}-\bar{y}_{i\bullet\bullet}-\bar{y}_{\bullet j\bullet}+\bar{y}_{\bullet\bullet\bullet})+(y_{ijk}-\bar{y}_{ij\bullet})\Big]^2 \\
&= \sum_{i=1}^{a} bn(\bar{y}_{i\bullet\bullet}-\bar{y}_{\bullet\bullet\bullet})^2 + \sum_{j=1}^{b} an(\bar{y}_{\bullet j\bullet}-\bar{y}_{\bullet\bullet\bullet})^2 \\
&\quad + \underbrace{\sum_{i=1}^{a}\sum_{j=1}^{b} n(\bar{y}_{ij\bullet}-\bar{y}_{i\bullet\bullet}-\bar{y}_{\bullet j\bullet}+\bar{y}_{\bullet\bullet\bullet})^2}_{\text{Interaction Sum Sq.}} + \underbrace{\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk}-\bar{y}_{ij\bullet})^2}_{\text{full model Res Sum Sq.}} + \text{ a cross-product term which is zero .}
\end{aligned}
$$

Not examinable

Interaction sum of squares

- The "no interaction" model **residual sum of squares** can itself be decomposed into two pieces:
  - the *full model* **residual sum of squares** and
  - whatever is left over, which is

$$\sum_{i=1}^{a}\sum_{j=1}^{b} n(\bar{y}_{ij\bullet} - \bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet j\bullet} + \bar{y}_{\bullet\bullet\bullet})^2 \,.$$

  - This is the **interaction sum of squares**.
- To see why, note that the (i, j)-th *estimated interaction effect* is

$$(\widehat{\alpha\gamma})_{ij} = \hat{\mu}_{ij} - \left[\hat{\mu} + \hat{\alpha}_i + \hat{\gamma}_j\right] = \bar{y}_{ij\bullet} - \left[\bar{y}_{\bullet\bullet\bullet} + (\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet}) + (\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet})\right] = \bar{y}_{ij\bullet} - \bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet j\bullet} + \bar{y}_{\bullet\bullet\bullet}\,,$$

  so the **interaction sum of squares** is precisely n times the sum of squares of the *estimated interaction effects*.

F-test for no interaction

- As usual, under the full model, the (random variable version of the) **residual sum of squares**

$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(Y_{ijk}-\bar{Y}_{ij\bullet})^2 \sim \sigma^2\chi^2_{ab(n-1)}\,,$$

  since

  - the total sample size is $N = abn$;
  - the number of groups is $g = ab$ so
  - the degrees of freedom for residuals is $N - g = abn - ab = ab(n-1)$.

- It turns out that *under the no interaction model* (the null hypothesis) the (random variable version of the) interaction sum of squares

$$\sum_{i=1}^{a}\sum_{j=1}^{b} n(\bar{Y}_{ij\bullet}-\bar{Y}_{i\bullet\bullet}-\bar{Y}_{\bullet j\bullet}+\bar{Y}_{\bullet\bullet\bullet})^2 \sim \sigma^2\chi^2_{(a-1)(b-1)}$$

  (we have already established there are $(a-1)(b-1)$ df for interactions).

- The F-statistic for testing the null hypothesis of no interactions is

$$\frac{\text{Interaction Mean Square}}{\text{(full model) Residual Mean Square}} = \frac{\text{(Int Sum Sq.)}/[(a-1)(b-1)]}{\text{(full model Residual Sum Sq.)}/[ab(n-1)]} \sim F_{(a-1)(b-1),ab(n-1)}$$

  if the null hypothesis of no interactions is true.

Back to our examples

Poisons and antidotes

Recall we have two factors, `poison` with 3 levels:

```
levels(poison.data$poison)
```

```
## [1] "I"   "II"  "III"
```

and `antidote` with 4 levels:

```
levels(poison.data$antidote)
```

```
## [1] "A" "B" "C" "D"
```

poison:antidote

- Given the two separate factors `poison` and `antidote`, the R object `poison:antidote` is a new factor which has as its levels every possible `poison:antidote` combination:

```
poison.data$poison:poison.data$antidote
```

```
##  [1] I:A   I:A   I:A   I:A   II:A  II:A  II:A  II:A  III:A
## [10] III:A III:A III:A I:B   I:B   I:B   I:B   II:B  II:B
## [19] II:B  II:B  III:B III:B III:B III:B I:C   I:C   I:C
## [28] I:C   II:C  II:C  II:C  II:C  III:C III:C III:C III:C
## [37] I:D   I:D   I:D   I:D   II:D  II:D  II:D  II:D  III:D
## [46] III:D III:D III:D
## 12 Levels: I:A I:B I:C I:D II:A II:B II:C II:D ... III:D
```

- This factor indicates the 12 groups in the "big 1-way ANOVA" where each `poison:antidote` combination is a different treatment;
    - there are 4 observations on each "treatment".

```
poison.data %>%
  group_by(poison, antidote) %>%
  count()
```

```
## # A tibble: 12 × 3
## # Groups:   poison, antidote [12]
##    poison antidote     n
##    <fct>  <fct>    <int>
##  1 I      A            4
##  2 I      B            4
##  3 I      C            4
##  4 I      D            4
##  5 II     A            4
##  6 II     B            4
##  7 II     C            4
##  8 II     D            4
##  9 III    A            4
## 10 III    B            4
## 11 III    C            4
## 12 III    D            4
```

- We could fit the full model (the "big 1-way ANOVA"):

```
a1 = aov(inv_survival ~ poison:antidote, data = poison.data)
anova(a1)
```

```
## Analysis of Variance Table
##
## Response: inv_survival
##                 Df Sum Sq Mean Sq F value    Pr(>F)
## poison:antidote 11 56.862  5.1693  21.531 1.289e-12 ***
## Residuals       36  8.643  0.2401
## ---
## Signif. codes:
## 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- There is a clear treatment effect
- Note there are 47 df in total:
  - 11 for treatments
  - 36 for residuals

A better approach

- A full, appropriate two-factor ANOVA table can be produced by using an appropriate formula:
    - fit `poison:antidote` *after* fitting the main effects:

```
a3 = aov(inv_survival ~ poison + antidote + poison:antidote,
         data = poison.data)
summary(a3)
```

```
##                  Df Sum Sq Mean Sq F value   Pr(>F)
## poison            2  34.88  17.439   72.64 2.31e-13 ***
## antidote          3  20.41   6.805   28.34 1.38e-09 ***
## poison:antidote   6   1.57   0.262    1.09    0.387
## Residuals        36   8.64   0.240
## ---
## Signif. codes:
## 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- The sum of squares for `poison:antidote` here only contains the variation explained by treatments *not already explained by main effects*.
- This provides a convenient *partition* of the "big 1-way ANOVA" **treatment sum of squares** (11 df in total) into 3 contributions:
    - 2 df for `poison` main effect;
    - 3 df for `antidote` (antidote) main effect;
    - 6 df for the interaction effect.

Alternate formula for two-factor ANOVA

- An *equivalent* formula which includes main effects followed by interactions is given as follows:

```
summary(aov(inv_survival ~ poison * antidote, data = poison.data))
```

```
##                 Df Sum Sq Mean Sq F value   Pr(>F)
## poison           2  34.88  17.439   72.64 2.31e-13 ***
## antidote         3  20.41   6.805   28.34 1.38e-09 ***
## poison:antidote  6   1.57   0.262    1.09    0.387
## Residuals       36   8.64   0.240
## ---
## Signif. codes:
## 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- This is what most people use, although the equivalent (but longer) form used initially (i.e. `poison+antidote+poison:antidote`) is perhaps more useful for beginners.
- The `poison` and `antidote` main effects **are significant** (i.e. there is a significant difference between the levels of each treatment group).

Paper planes

- Let us perform a similar analysis on the Paper Planes data.
- First, we check the "big 1-way ANOVA":

```
summary(aov(Distance ~ Paper:Plane, data = planes))
```

```
##             Df   Sum Sq Mean Sq F value  Pr(>F)
## Paper:Plane  3 25491258 8497086   10.66 0.00106 **
## Residuals   12  9561029  796752
## ---
## Signif. codes:
## 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

There is a strong Treatment effect.

- Now we partition the **treatment sum of squares** into main effects and interactions:

```
plane_aov = aov(Distance ~ Paper * Plane, data = planes)
summary(plane_aov)
```

```
##             Df   Sum Sq  Mean Sq F value   Pr(>F)
## Paper        1  1718721  1718721   2.157 0.167628
## Plane        1   385641   385641   0.484 0.499861
## Paper:Plane  1 23386896 23386896  29.353 0.000156 ***
## Residuals   12  9561029   796752
## ---
## Signif. codes:
## 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- There is a *strong* interaction effect.
- Note the large p-values for the main effects.
  - Are they significant? No.
  - Should they be dropped from the model: **absolutely not**!
- If an interaction is significant, we must retain the **full model**.

Interaction plots

Interaction plots

- We can graphically examine interactions using **interaction plots**.
- These involve choosing one of the factors as the `x.factor` and one as the `trace.factor`:
  - levels of the `x.factor` are marked, equally spaced, on the "x-axis";
  - then a *trace* of each level of the (other) `trace.factor` is created by
    - plotting mean responses for that level against the corresponding level on the x-axis;
    - joining the mean responses within in each level by a piecewise linear curve.
- If there is **no interaction** the traces should be "roughly parallel".
- If there **is** an interaction, the traces may cross or deviate from parallelism in some other way.

The notation `x.factor` and `trace.factor` come from the R function `interaction.plot()` but we will use **ggplot2**.

Poisons/antidotes interaction plots

- The mean responses by treatment combination:

```
sum_dat = poison.data %>% group_by(poison, antidote) %>%
  summarise(mean = mean(inv_survival))
glimpse(sum_dat)
```

```
## Rows: 12
## Columns: 3
## Groups: poison [3]
## $ poison   <fct> I, I, I, I, II, II, II, II, III, III, III…
## $ antidote <fct> A, B, C, D, A, B, C, D, A, B, C, D
## $ mean     <dbl> 2.486881, 1.163464, 1.862724, 1.689682, 3…
```

```
sum_dat %>% spread(key = poison, value = mean)
```

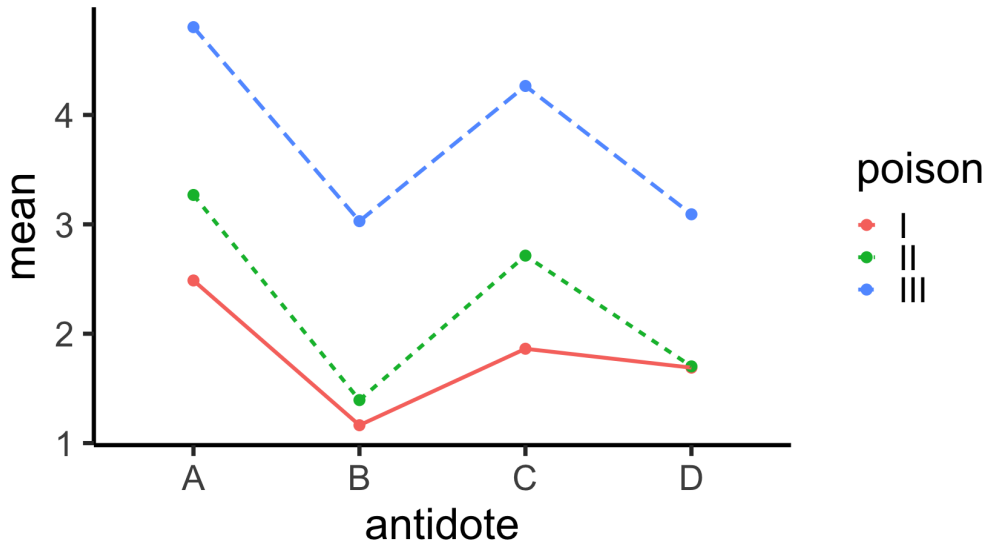
```
## # A tibble: 4 × 4
##   antidote     I    II   III
##   <fct>    <dbl> <dbl> <dbl>
## 1 A         2.49  3.27  4.80
## 2 B         1.16  1.39  3.03
## 3 C         1.86  2.71  4.26
```

wide

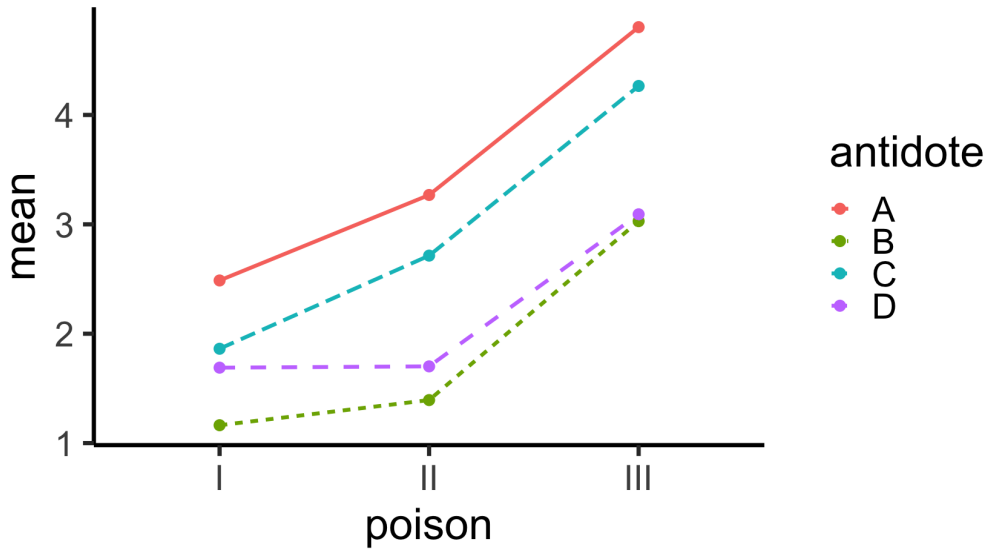
id	x	y	z
1	a	c	e
2	b	d	f

Poisons/antidotes interaction plots

```
sum_dat %>%
  ggplot(aes(x = antidote, y = mean,
             group = poison,
             linetype = poison,
             colour = poison)) +
  geom_point(size = 4) +
  geom_line(size = 1.5) +
  theme_classic(base_size = 36)
```



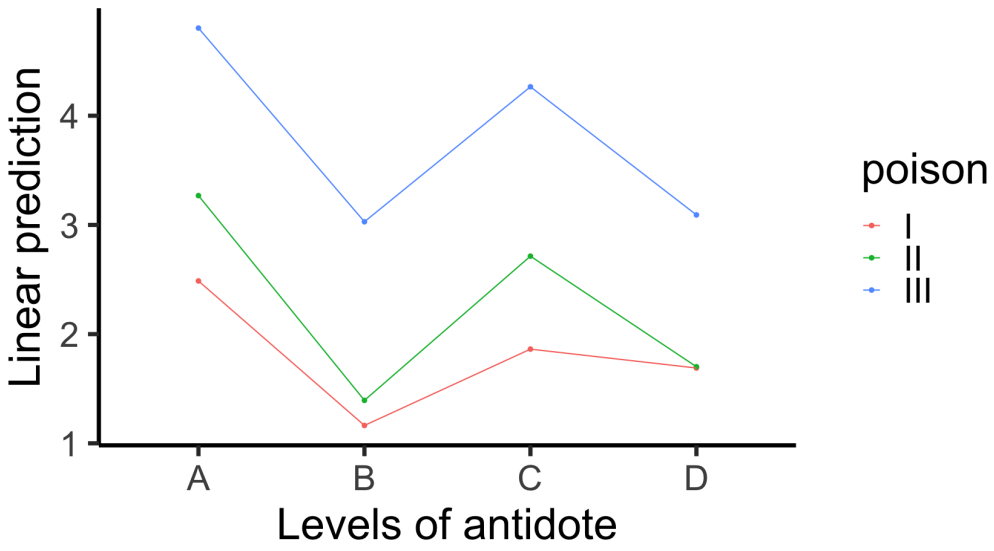

```
sum_dat %>%
  ggplot(aes(x = poison, y = mean,
             group = antidote,
             linetype = antidote,
             colour = antidote)) +
  geom_point(size = 4) +
  geom_line(size = 1.5) +
  theme_classic(base_size = 36)
```



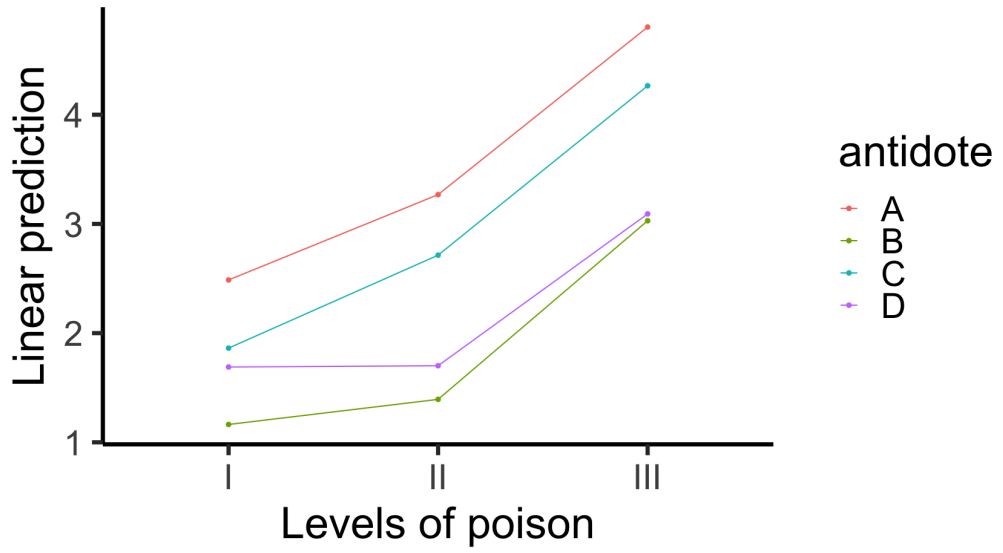
Using emmeans

The `emmip()` function from the **emmeans** package can do this for us:

```
library(emmeans)
emmip(a1, poison ~ antidote) +
  theme_classic(base_size = 36)
```




```
emmip(a1, antidote ~ poison) +
  theme_classic(base_size = 36)
```



`emmip()` is named after: estimated marginal means interaction plot.

Paper planes interaction plots

- Recall that there was a strong interaction effect.
- What would be expect to see here?

```
plane_sum_dat = planes %>%
  group_by(Plane, Paper) %>%
  summarise(mean = mean(Distance))
plane_sum_dat
```

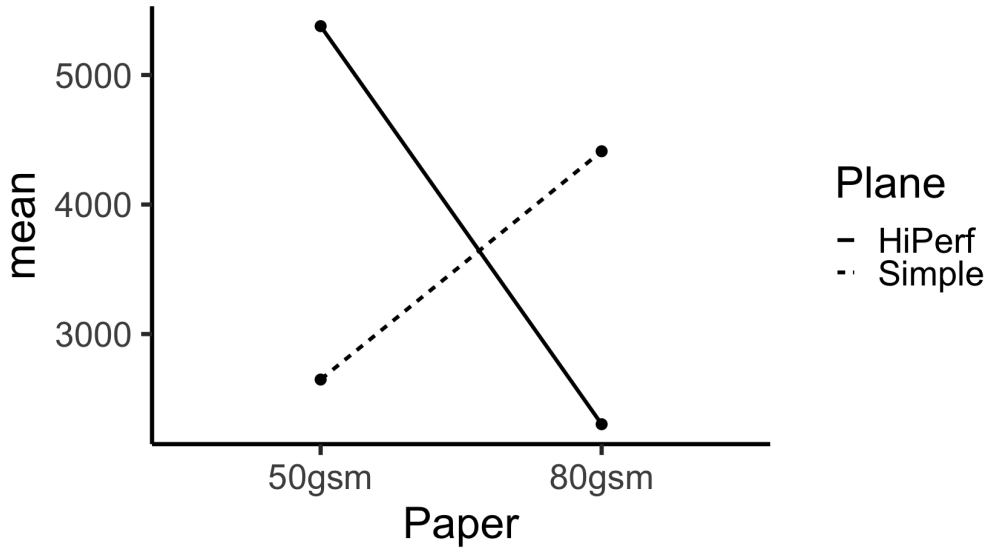
```
## # A tibble: 4 × 3
## # Groups:   Plane [2]
##   Plane  Paper  mean
##   <chr>  <chr> <dbl>
## 1 HiPerf 50gsm 5377.
## 2 HiPerf 80gsm 2304.
## 3 Simple 50gsm 2649.
## 4 Simple 80gsm 4411.
```

```
plane_sum_dat %>%
  spread(key = Plane, value = mean)
```

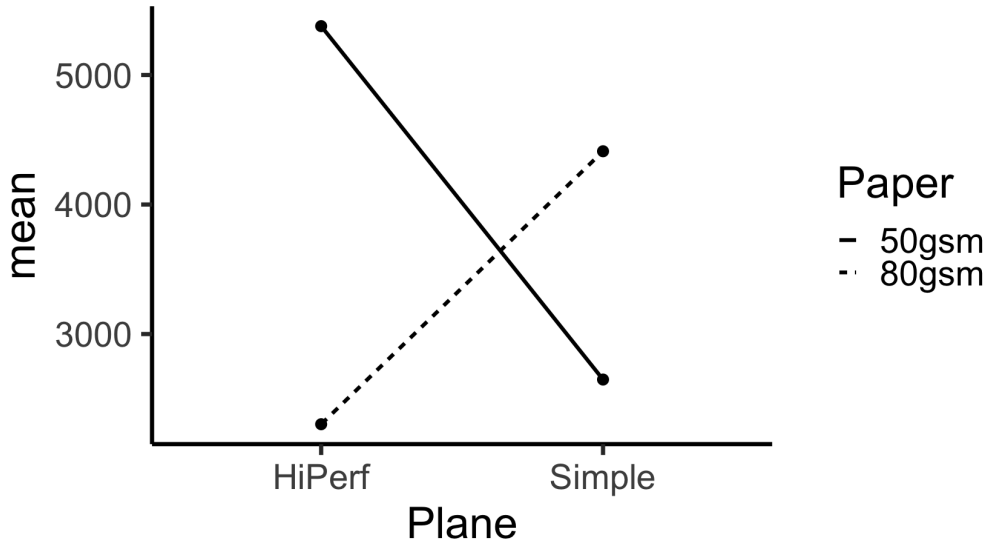
```
## # A tibble: 2 × 3
##   Paper HiPerf Simple
##   <chr>  <dbl>  <dbl>
## 1 50gsm  5377.  2649.
## 2 80gsm  2304.  4411.
```

Paper planes interaction plots

```
plane_sum_dat %>%
  ggplot(aes(x = Paper, y = mean,
             group = Plane,
             linetype = Plane)) +
  geom_point(size = 4) +
  geom_line(size = 1.5) +
  theme_classic(base_size = 36)
```



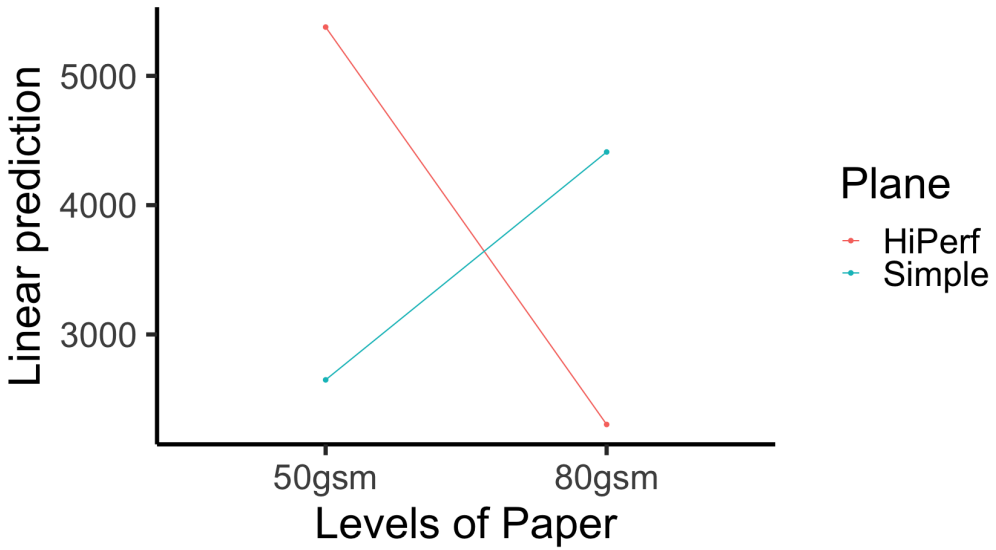

```
plane_sum_dat %>%
  ggplot(aes(x = Plane, y = mean,
             group = Paper,
             linetype = Paper)) +
  geom_point(size = 4) +
  geom_line(size = 1.5) +
  theme_classic(base_size = 36)
```



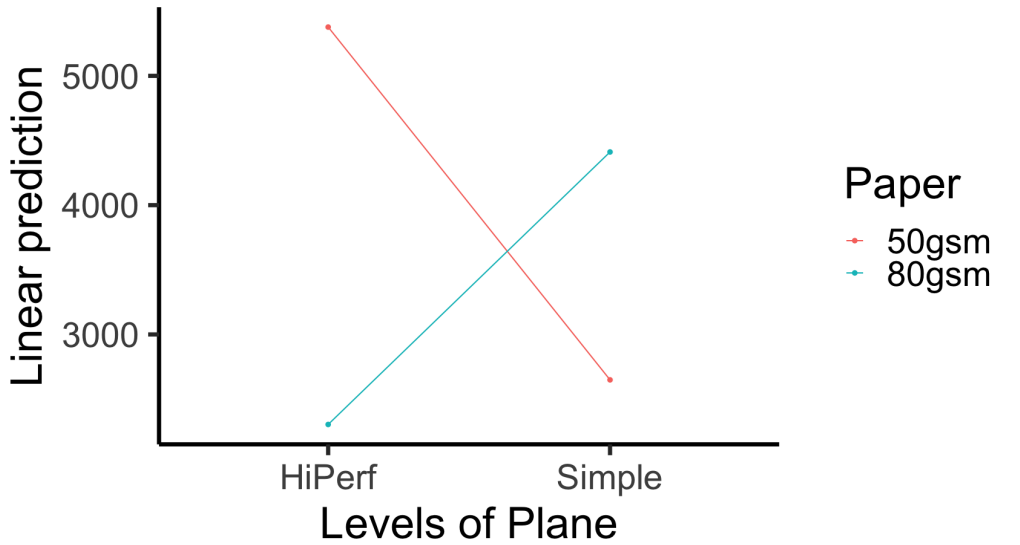
Using emmeans

The `emmip()` function from the **emmeans** package can do this for us:

```
emmip(plane_aov, Plane ~ Paper) +
  theme_classic(base_size = 36)
```




```
emmip(plane_aov, Paper ~ Plane) +
  theme_classic(base_size = 36)
```



Crossing traces: interaction

- The traces cross dramatically, highlighting the **strong interaction effect**.
- What is going on here?
  - The high performance design flies further with lighter paper;
  - the simple design flies further with heavier paper.

Post hoc comparisons

Post hoc comparisons

- As with any analysis of variance, once we establish that there is something going on, we would like to "zoom in" and learn more about what is leading to the significance.
- **However** in this two-factor scenario, *exactly which* comparisons one might be interested in *may depend on whether there was a significant interaction or not*.
- If there **is a significant interaction**, we might like to know if one factor has a strong effect of the *other* factor within some or all of its levels.
- If there is **no significant interaction**, the levels of factor A can be compared "independently" of levels of factor B. We may wish to determine:
  - which levels of factor A are significantly different from one another, or
  - which levels of factor B are significantly different from one another.
- The main thing to realise is that any within-factor comparison is, at the end of the day, just a *contrast* in the original μ_{ij}'s and so an appropriate multiple comparisons adjustment can be made accordingly.

Comparing poisons

- We determined that there is no interaction between `poison` and `antidote`, so we can perhaps go ahead and compare different `poison` (to each other) and also compare antidote treatments (`antidote`) (to each other).

```
poison.data %>%
  ggplot(aes(x = poison, y = inv_survival)) +
  geom_boxplot() +
  theme_classic(base_size = 30)
```

```
poison.data %>%
  ggplot(aes(x = antidote, y = inv_survival)) +
  geom_boxplot() +
  theme_classic(base_size = 30)
```

Pairwise difference t-statistics

```
a2 = aov(inv_survival ~ poison + antidote, data = poison.data)
emmeans(a2, ~ poison) %>% contrast(method = "pairwise", adjust = "none")
```

```
##  contrast estimate    SE df t.ratio p.value
##  I - II     -0.469 0.174 42  -2.688  0.0103
##  I - III    -1.996 0.174 42 -11.451  <.0001
##  II - III   -1.528 0.174 42  -8.763  <.0001
##
## Results are averaged over the levels of: antidote
```

```
emmeans(a2, ~ antidote) %>% contrast(method = "pairwise", adjust = "none")
```

```
##  contrast estimate    SE df t.ratio p.value
##  A - B       1.657 0.201 42   8.233  <.0001
##  A - C       0.572 0.201 42   2.842  0.0069
##  A - D       1.358 0.201 42   6.747  <.0001
##  B - C      -1.085 0.201 42  -5.391  <.0001
##  B - D      -0.299 0.201 42  -1.485  0.1449
##  C - D       0.786 0.201 42   3.905  0.0003
##
## Results are averaged over the levels of: poison
```

Assessing significance

- We have a few options, when it comes to comparing these:
  - use a Bonferroni approach;
  - Tukey's method;
  - Scheffe's method.
- There may be some doubt as to the validity of the first two, depending on whether the decision of **which** comparisons to look at came **after the initial hypothesis test** or not.
- **If** the 3+6=9 within-factor comparisons were planned *before looking at the data* then a Bonferroni adjustment could be made as follows:
  - multiply all *unadjusted* two-sided t-test p-values by 9.

```
p_emm = contrast(emmeans(a2, ~poison), method = "pairwise", adjust = "bonferroni")
p_emm
```

```
##  contrast estimate    SE df t.ratio p.value
##  I - II     -0.469 0.174 42  -2.688  0.0308
##  I - III    -1.996 0.174 42 -11.451  <.0001
##  II - III   -1.528 0.174 42  -8.763  <.0001
##
## Results are averaged over the levels of: antidote
## P value adjustment: bonferroni method for 3 tests
```

```
a_emm = contrast(emmeans(a2, ~antidote), method = "pairwise", adjust = "bonferroni")
a_emm
```

```
##  contrast estimate    SE df t.ratio p.value
##  A - B       1.657 0.201 42   8.233  <.0001
##  A - C       0.572 0.201 42   2.842  0.0414
##  A - D       1.358 0.201 42   6.747  <.0001
##  B - C      -1.085 0.201 42  -5.391  <.0001
##  B - D      -0.299 0.201 42  -1.485  0.8693
##  C - D       0.786 0.201 42   3.905  0.0020
##
## Results are averaged over the levels of: poison
## P value adjustment: bonferroni method for 6 tests
```

Bonferroni method

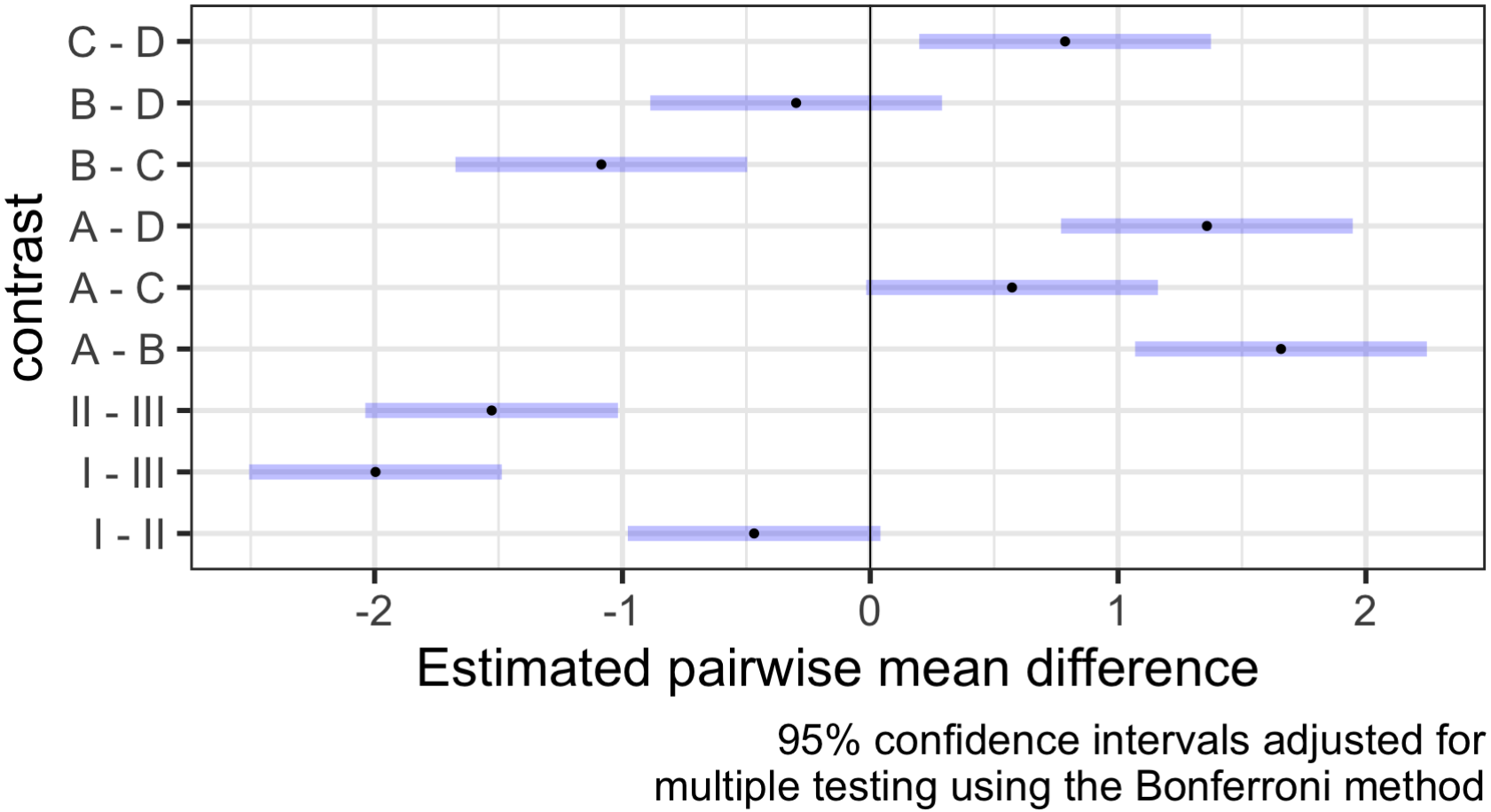
The **emmeans** package didn't know we were looking at **9** tests. But we can tell it that we are:

```
pa_emm = update(p_emm + a_emm)
pa_emm
```

```
##  contrast estimate    SE df t.ratio p.value
##  I - II     -0.469 0.174 42  -2.688  0.0924
##  I - III    -1.996 0.174 42 -11.451  <.0001
##  II - III   -1.528 0.174 42  -8.763  <.0001
##  A - B       1.657 0.201 42   8.233  <.0001
##  A - C       0.572 0.201 42   2.842  0.0620
##  A - D       1.358 0.201 42   6.747  <.0001
##  B - C      -1.085 0.201 42  -5.391  <.0001
##  B - D      -0.299 0.201 42  -1.485  1.0000
##  C - D       0.786 0.201 42   3.905  0.0030
##
## Results are averaged over some or all of the levels of: antidote, poison
## P value adjustment: bonferroni method for 9 tests
```

- For `poison`: the `I-III` and `II-III` differences are highly significant, while the `I-II` has an adjusted p-value of about 0.09 (not really significant).
- For `antidote`: all differences are highly significant except `B-D` (not at all significant) and `A-C` (p=0.06).

```
plot(pa_emm) + theme_bw(base_size = 30) + geom_vline(xintercept = 0) +
  labs(x = "Estimated pairwise mean difference",
       caption = "95% confidence intervals adjusted for\nmultiple testing using the Bonferroni metho
```



Paper planes

- For this case, where the full model must be retained, we really do have a "big 1-way ANOVA".
- In such a case, all pairwise comparisons between *all treatment combinations*, or some other set of comparisons, may be of interest. For example, we can use Tukey's all pairwise comparisons:

```
plane_aov = aov(Distance ~ Plane*Paper, data = planes)
plane_emm = emmeans(plane_aov,  ~ Paper + Plane)
contrast(plane_emm, method = "pairwise", adjust = "tukey")
```

```
##  contrast                    estimate  SE df t.ratio p.value
##  50gsm HiPerf - 80gsm HiPerf     3074 631 12   4.870  0.0019
##  50gsm HiPerf - 50gsm Simple     2728 631 12   4.323  0.0047
##  50gsm HiPerf - 80gsm Simple      966 631 12   1.530  0.4508
##  80gsm HiPerf - 50gsm Simple     -345 631 12  -0.547  0.9457
##  80gsm HiPerf - 80gsm Simple    -2108 631 12  -3.339  0.0262
##  50gsm Simple - 80gsm Simple    -1762 631 12  -2.792  0.0677
##
## P value adjustment: tukey method for comparing a family of 4 estimates
```

- Note that all *within-level* comparisons are quite significant.

Want to know more?

Applied Linear Models - STAT3022 runs in Semester 1.

References

Barrios, E. (2016). *BHH2: Useful Functions for Box, Hunter and Hunter II*. R package version 2016.05.31. URL: https://CRAN.R-project.org/package=BHH2.

Box, G. E. P. and D. R. Cox (1964). "An Analysis of Transformations". In: *Journal of the Royal Statistical Society. Series B (Methodological)* 26.2, pp. 211-252. ISSN: 00359246. URL: http://www.jstor.org/stable/2984418.

Lenth, R. (2018). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. R package version 1.2.3. URL: https://CRAN.R-project.org/package=emmeans.

Mackisack, M. (1994). "What Is the Use of Experiments Conducted By Statistics Students?" In: *Journal of Statistics Education* 2.1. URL: http://www.amstat.org/publications/jse/v2n1/mackisack.html.