

Lab 5: Post-hoc Tests & Factorial ANOVA

PSC 103B

Marwin Carmo

Today's dataset

- We'll be using a modified version of last week's data set.
- I randomly chose 68 penguins from each species, because the *manual* post-hoc test we are doing requires equal group sizes.

```
1 penguins_data <- read.csv("data/penguins_103b.csv")
```



Post-Hoc Tests

- The one-way ANOVA we conducted last week we were interested in whether the different species of penguins had, on average, equal bill lengths, or whether there was some difference in the average bill length among the 3 species.

```
1 my_anova <- aov(bill_length_mm ~ species, data = penguins_data)
2 summary(my_anova)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
species	2	4436	2217.9	241.8	<2e-16 ***
Residuals	201	1844	9.2		

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

- What can we conclude?

Post-Hoc Tests

- After a significant ANOVA result, we're interested in learning *where* those differences are.
- Post-hoc tests are meant to correct for the many comparisons you are doing, so that your chances of a **Type 1 error** (saying 2 groups have different means when they actually have the same means) stays at .05.
- There are multiple post-hoc tests available, and they differ in how conservative they are. We're focus on **Tukey's Honest Significant Difference (HSD) Test**.

Tukey's HSD

- It computes a value called a “**Critical Difference**”, and so any group differences that are larger than the Critical Difference (CD) in absolute value are significant.

$$CD = q \times \sqrt{\frac{MS_w}{n}}$$

- where n is the group size within each group, and MS_w is the Mean Squared Within you can get from your ANOVA table.

Find n and MS_w

n

```
1 table(penguins_data$species)
```

```
Adelie Chinstrap   Gentoo  
68      68      68
```

```
1 n <- 68
```

MS_w (tip: look in the ANOVA summary)

```
1 summary(my_anova)
```

```
      Df Sum Sq Mean Sq F value Pr(>F)  
species    2  4436   2217.9    241.8 <2e-16 ***  
Residuals 201  1844     9.2  
---
```

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

```
1 MSW <- 9.2
```

To find q , we can use the function `qtukey()`, that takes 3 main arguments:

- **p**, which is the probability to the *left* of the critical value,
- **nmeans**, which is the number of groups we're comparing and,
- **df**, which is just df within
- If our test is conducted at the .05 level, then the probability *below* the critical value is .95.

Exercise

Find q using the `qtukey()` function. Set the values for the arguments **p**, **nmeans**, and **df**. Save your result to an object called `qvalue`. Tip: you can also find df within in the ANOVA summary

Find q

```
1 qvalue <- qtkey(p = .95,  
2   nmeans = 3, # because we have 3 species  
3   df = 201)  
4 qvalue
```

```
[1] 3.339252
```



Tukey's HSD

- Now, we can go ahead and calculate our CD.

```
1 cd <- qvalue * sqrt(MSW / n)
2 cd
```

```
[1] 1.228255
```

- How can we get all the groups' means?

```
1 penguin_means <- tapply(penguins_data$bill_length_mm,
2   penguins_data$species,
3   mean, na.rm = TRUE)
4 penguin_means
```

```
Adelie Chinstrap Gentoo
38.38529 48.83382 47.60588
```

Exercise

Find all the possible mean differences between groups and compare to the critical difference we just calculated. What can we conclude about those differences?

```
1 adlie_mean <- penguin_means[1]
2 chinstrap_mean <- penguin_means[2]
3 gentoo_mean <- penguin_means[3]
```

```
1 adlie_mean - chinstrap_mean
```

Adelie
-10.44853

```
1 adlie_mean - gentoo_mean
```

Adelie
-9.220588

```
1 chinstrap_mean - gentoo_mean
```

Chinstrap
1.227941

Tukey's HSD

Comparison	Mean difference
Adelie - Chinstrap	-10.45
Adelie - Gentoo	-9.22
Chinstrap - Gentoo	1.23

- Given our CD of 1.21 which of these differences are significant?
- We can say that Adelie has a significantly lower average bill length than both Chinstrap and Gentoo.
- And Chinstrap has a significantly longer bill length, on average, than Gentoo.

Tukey's HSD

R has a built-in function for Tukey's HSD, called `TukeyHSD()`, and the only argument you need to give it is the ANOVA object you saved earlier:

```
1 TukeyHSD(my_anova)
```

Tukey multiple comparisons of means
95% family-wise confidence level

Fit: aov(formula = bill_length_mm ~ species, data = penguins_data)

\$species	diff	lwr	upr	p adj
Chinstrap-Adelie	10.448529	9.222071	11.674988263	0.0000000
Gentoo-Adelie	9.220588	7.994129	10.447047086	0.0000000
Gentoo-Chinstrap	-1.227941	-2.454400	-0.001482326	0.0496478

Factorial ANOVA

Factorial ANOVA

Factorial ANOVA allow us examine in one test:

- is there an effect of Grouping Variable 1 (at least one mean is different in those groups)?
- is there an effect of Grouping Variable 2?
- is there an interaction between the 2 variables (so that the effect of Grouping Variable 1 depends on Grouping Variable 2, or vice versa)?

Factorial ANOVA

- Let's assume we were interested in whether **body mass** differed not only among the 3 penguin **species**, but also whether it differed among **males and females**.
- If we were to do this as 2 separate ANOVAs, we would have 2 sets of hypotheses:
 - One for the species of penguin:
 - $H_0 : \mu_{Adelie} = \mu_{Chinstrap} = \mu_{Gentoo}$ (there is no difference in average body mass among the species).
 - H_A : At least one mean is significantly different from the rest.

Factorial ANOVA

- One for the sex:
 - $H_0 : \mu_{female} = \mu_{male}$ (there is no difference in average body mass among male and female penguins).
 - $H_A : \mu_{female} \neq \mu_{male}$ (note that because there are only 2 groups, we can be more specific with our alternative hypothesis).
- And we also have a hypothesis for the interaction:
 - H_0 : There is no interaction between the species of penguin and their sex.
 - H_A : There is an interaction between the species of penguin and their sex.

Factorial ANOVA

- In R where we “multiply” our two predictors to include an interaction between them.

```
1 factorial_anova <- aov(body_mass_g ~ species * sex, data = penguins_data)
2 summary(factorial_anova)
```

```
      Df Sum Sq Mean Sq F value    Pr(>F)
species  2  93491533  46745767  466.899 < 2e-16 ***
sex       1  19697662  19697662  196.741 < 2e-16 ***
species:sex  2  1480819   740409    7.395 0.000801 ***
Residuals 196 19623451  100120
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
2 observations deleted due to missingness
```

- Which results are significant? What does this tell us?

Factorial ANOVA

Main effect of species

At least one species has an average body mass that is significantly different from another species.

1 TukeyHSD(factorial_anova)\$species

	diff	lwr	upr	p adj
Chinstrap-Adelie	81.5957	-47.03703	210.2284	2.940693e-01
Gentoo-Adelie	1484.3284	1355.22009	1613.4366	7.571721e-14
Gentoo-Chinstrap	1402.7327	1274.09993	1531.3654	7.571721e-14

Factorial ANOVA

Main effect of sex

The p -value is also less than .05 so we can reject H_0 and (since there are only 2 groups) say that there is a significant difference between the body mass of male and female penguins

```
1 TukeyHSD(factorial_anova)$sex
```

	diff	lwr	upr	p adj
male-female	620.0387	532.1883	707.8891	7.571721e-14

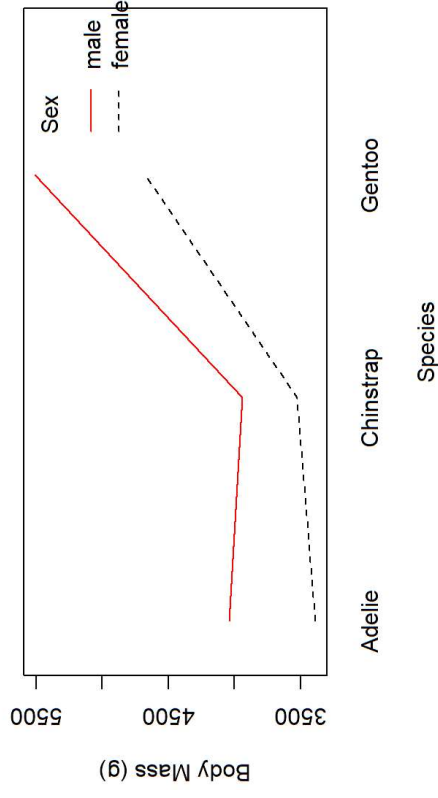
Factorial ANOVA

Interaction

- The difference in body mass between male and female penguins depends on the species.
- We can try to uncover what that interaction is by graphing the means.
- In this graph, we would have one grouping variable on the x-axis and the other grouping variable as different lines on the graph.
- To do this, we will use the function `interaction.plot()`.

Factorial ANOVA

```
1 interaction.plot(x.factor = penguins_data$species, # grouping variable on x axis
2                 trace.factor = penguins_data$sex, # grouping variable as lines
3                 response = penguins_data$body_mass_g, # outcome variable
4                 fun = mean, # summary statistic
5                 type = "l", # graph lines
6                 col = c("black", "red"), xlab = "Species", ylab = "Body Mass (g)", trace.label = "Sex")
```



Factorial ANOVA

- Males, on average, weigh more than females.
- Gentoo penguins weigh more than Adelie and Chinstrap, but Adelie and Chinstrap are pretty close to each other.
- The interaction appears to be that the magnitude of the difference is different depending on the species. In particular, males and females have a much smaller difference in the Chinstrap species than the other 2 species.

Now you try

Exercise

Run a factorial ANOVA testing the interaction between the penguin `sex` and its `island` in accounting for differences in `body_mass_g`. What can you conclude?

```
1 factorial_anova_2 <- aov(body_mass_g ~ sex * island, data = penguins_data)
2 summary(factorial_anova_2)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
sex	1	28677875	28677875	106.026	< 2e-16 ***
island	2	49789988	24894994	92.040	< 2e-16 ***
sex:island	2	2811759	1405880	5.198	0.00632 **
Residuals	196	53013843	270479		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
2 observations deleted due to missingness

Now you try

At least penguins from one island have an average body mass that is significantly different from penguins from another island.

```
1 TukeyHSD(factorial_anova_2)$island
```

	diff	lwr	upr	p adj
Dream-Biscoe	-1000.52117	-1183.1127	-817.9296	7.571721e-14
Torgersen-Biscoe	-989.77563	-1287.4312	-692.1201	8.849588e-13
Torgersen-Dream	10.74554	-286.6004	308.0915	9.959924e-01

Males have significantly higher body mass than females.

```
1 TukeyHSD(factorial_anova_2)$sex
```

	diff	lwr	upr	p adj
male-female	753.9099	609.5154	898.3044	9.325873e-14

