

Midterm2 Review

Nate Wells

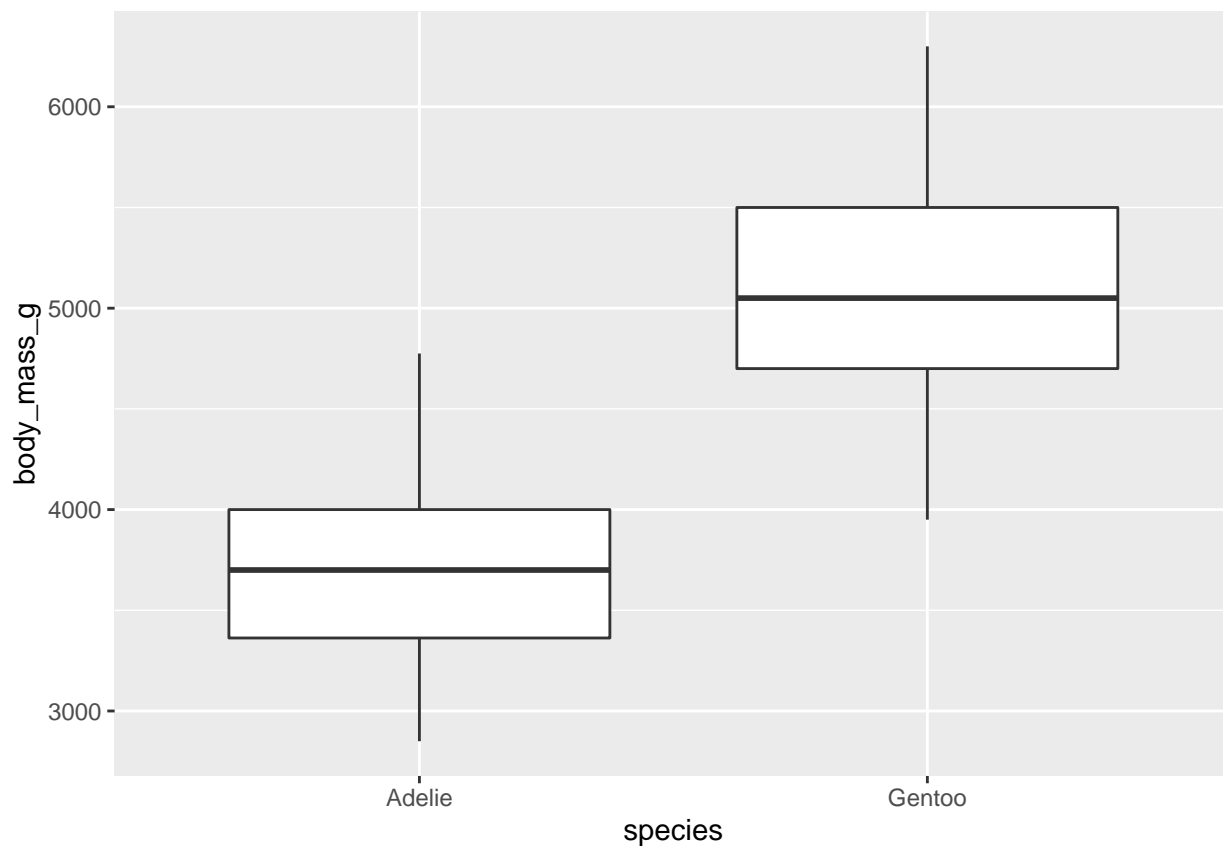
4/2/2021

```
penguins <- penguins %>% drop_na()
```

Can we distinguish between Adelie and Gentoo penguin based on bodymass?

```
penguins_ag <- penguins %>% filter(species %in% c("Adelie", "Gentoo"))
```

```
penguins_ag %>% ggplot(aes(x = species, y = body_mass_g))+geom_boxplot()
```



Perform hypothesis test to determine whether body mass for two species is equal

Case I

Null is

$$H_0 : \mu_a = \mu_g$$

Alt Hyp is

$$H_a : \mu_a \neq \mu_g$$

Case II

Null is

$$H_0 : \mu_a = \mu_g$$

Alt Hyp is

$$H_a : \mu_a < \mu_g$$

Case III

Null is

$$H_0 : \mu_a = \mu_g$$

Alt Hyp is

$$H_a : \mu_a > \mu_g$$

```
penguin_stats<-penguins_ag %>%  
  specify(response = body_mass_g, explanatory = species ) %>%  
  hypothesise(null = "independence") %>%  
  generate(reps = 2000, type= "permute") %>%  
  calculate(stat ="diff in means", order = c("Gentoo", "Adelie"))
```

```
og_stat <-penguins_ag %>%  
  specify(response = body_mass_g, explanatory = species ) %>%  
  calculate(stat ="diff in means", order = c("Gentoo", "Adelie"))
```

```
penguin_stats %>% get_p_value(obs_stat = og_stat, direction = "both")
```

```
## Warning: Please be cautious in reporting a p-value of 0. This result is an  
## approximation based on the number of `reps` chosen in the `generate()` step. See  
## `?get_p_value()` for more information.
```

```
## # A tibble: 1 x 1  
##   p_value  
##   <dbl>  
## 1      0
```

Compare p-value to significance level α . Standard choice $\alpha = 0.05$. In this case, $P\text{-value} < .05$, we reject H_0 in favor of the H_a .