Week 8 Lab

AUTHOR
Jessica Tran

PUBLISHED
April 5, 2024

Packages

We will need the following packages for this lab:

```
        library("car")
```

```
Loading required package: carData
```

```
        library("palmerpenguins")
```

Data

For this lab, we will be using the data set `penguins`. Take some time to get familiar with the data set using the help function.

```
        ?penguins
```

Problem 1

Using the data set `penguins`, perform a Two Way ANOVA test to see if there is any significant difference between the average `body_mass_g` of the penguins based off the factor `species` and the group `sex`. Use a significance level of 0.05

First, check that `species` and `sex` are both factors using the `head` command. If necessary, convert your data to factors.

```
        head(penguins)
```

```
# A tibble: 6 × 8
  species island    bill_length_mm bill_depth_mm flipper_length_mm body_mass_g
  <fct>   <fct>              <dbl>         <dbl>             <int>       <int>
1 Adelie  Torgersen           39.1          18.7               181        3750
2 Adelie  Torgersen           39.5          17.4               186        3800
3 Adelie  Torgersen           40.3          18                 195        3250
4 Adelie  Torgersen           NA            NA                  NA          NA
5 Adelie  Torgersen           36.7          19.3               193        3450
```

```
6 Adelie  Torgersen           39.3          20.6               190        3650
# ℹ 2 more variables: sex <fct>, year <int>
```

Now, conduct the test.

```
output <- aov(body_mass_g ~ species + sex, data = penguins)
```

Check the normality assumption.

```
shapiro.test(output$residuals)
```

```
    Shapiro-Wilk normality test

data:  output$residuals
W = 0.99734, p-value = 0.869
```

Check the equal variance assumption.

```
leveneTest(body_mass_g ~ species * sex, data = penguins)
```

```
Levene's Test for Homogeneity of Variance (center = median)
       Df F value Pr(>F)
group   5  1.3908 0.2272
      327
```

Check the balanced assumption.

```
table(penguins$species, penguins$sex)
```

```
           female male
  Adelie       73   73
  Chinstrap    34   34
  Gentoo       58   61
```

Display the results.

```
summary(output)
```

```
             Df    Sum Sq  Mean Sq F value Pr(>F)
species       2 145190219 72595110   724.2 <2e-16 ***
sex           1  37090262 37090262   370.0 <2e-16 ***
Residuals   329  32979185   100241
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
11 observations deleted due to missingness
```

If appropriate, perform a post hoc analysis.

```
TukeyHSD(output)
```

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

Fit: aov(formula = body_mass_g ~ species + sex, data = penguins)

$species
                       diff        lwr      upr     p adj
Chinstrap-Adelie   26.92385  -82.51532  136.363 0.8313289
Gentoo-Adelie    1386.27259 1294.21284 1478.332 0.0000000
Gentoo-Chinstrap 1359.34874 1246.03315 1472.664 0.0000000

$sex
                diff     lwr      upr p adj
male-female 667.4577 599.193 735.7224     0
```

Problem 2

Using the data set `penguins`, perform a Three Way ANOVA test to see if there is any significant difference between the average `body_mass_g` of the penguins based off the factor `species` and the group `sex`, and to determine if there is an interaction. Use a significance level of 0.05.

```
output1 <- aov(body_mass_g ~ species * sex, data = penguins)
summary(output1)
```

```
             Df    Sum Sq  Mean Sq F value   Pr(>F)
species       2 145190219 72595110 758.358  < 2e-16 ***
sex           1  37090262 37090262 387.460  < 2e-16 ***
species:sex   2   1676557   838278   8.757 0.000197 ***
Residuals   327  31302628    95727
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
11 observations deleted due to missingness
```

Submitting

Submit the following to Canvas:

- Your rendered PDF titled Lastname_8R. Make sure your name is at the top of the document.
- Your .qmd file