R code: F tests for ANOVA (Sleuth3 Section 5.3)

We will refine this R code next class

2019-09-20

Iris Flowers Example

Conduct a claim of the test that the population mean sepal width is the same for all three species.

$H_0 : \mu_1 = \mu_2 = \mu_3$

H_A : it is not the case that all three means are equal.

Any approach will start off by fitting the full model

```r
full_fit <- lm(Sepal.Width ~ Species, data = iris)
```

Option 1: F statistic and p-value at bottom of model summary

```r
summary(full_fit)
```

```
##
## Call:
## lm(formula = Sepal.Width ~ Species, data = iris)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.128 -0.228  0.026  0.226  0.972
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        3.42800    0.04804  71.359  < 2e-16 ***
## Speciesversicolor -0.65800    0.06794  -9.685  < 2e-16 ***
## Speciesvirginica  -0.45400    0.06794  -6.683 4.54e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3397 on 147 degrees of freedom
## Multiple R-squared:  0.4008, Adjusted R-squared:  0.3926
## F-statistic: 49.16 on 2 and 147 DF,  p-value: < 2.2e-16
```

Option 2: Call anova on your model fit

```r
anova(full_fit)
```

```
## Analysis of Variance Table
##
## Response: Sepal.Width
##            Df Sum Sq Mean Sq F value    Pr(>F)
## Species     2 11.345  5.6725   49.16 < 2.2e-16 ***
## Residuals 147 16.962  0.1154
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Option 3: Fit a reduced model and call `anova`, supplying both the full and reduced models.

For the reduced model, our explanatory variable is "1" – which means, include only an intercept.

```
reduced_fit <- lm(Sepal.Width ~ 1, data = iris)
anova(reduced_fit, full_fit)
```

```
## Analysis of Variance Table
##
## Model 1: Sepal.Width ~ 1
## Model 2: Sepal.Width ~ Species
##   Res.Df    RSS Df Sum of Sq     F    Pr(>F)
## 1    149 28.307
## 2    147 16.962  2    11.345 49.16 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

This third approach is too much work for the current example, but it is necessary in the examples we will look at next week.