

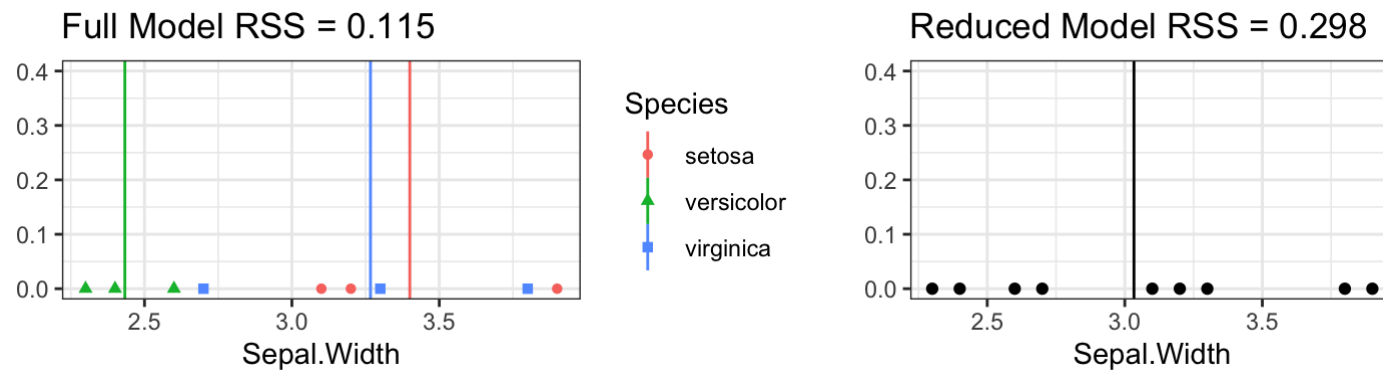
F Test p-values

Evan L. Ray

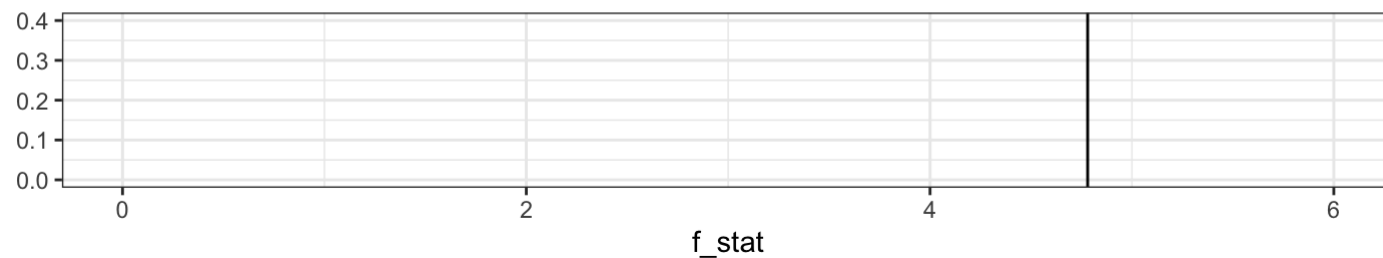
September 23, 2019

F tests: Iris example

Suppose we have 3 flowers of each iris species, and we want to conduct an F test of $H_0 : \mu_1 = \mu_2 = \mu_3$.



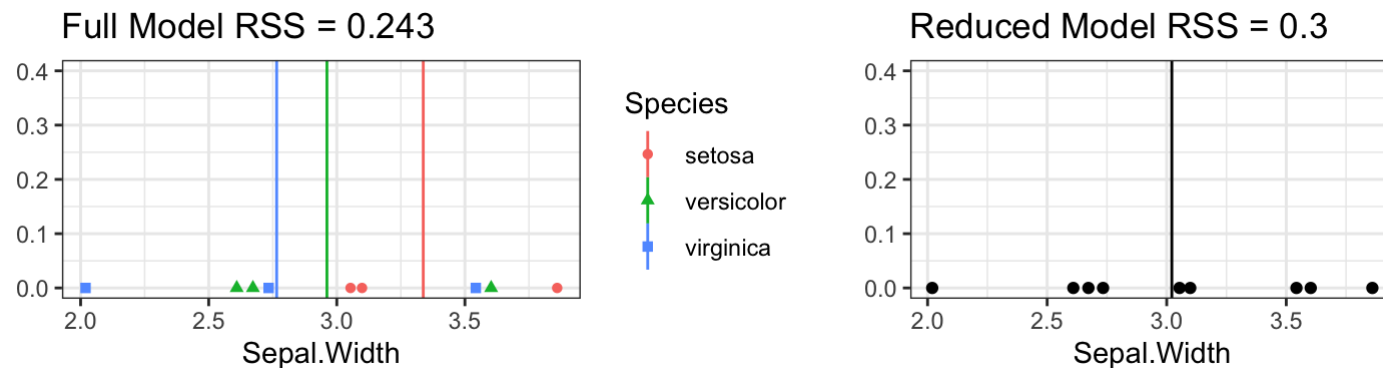
$$F = \frac{(\text{Extra Sum of Squares})/(\text{Extra Degrees of Freedom})}{(\text{Residual Sum of Squares, Full Model})/(\text{Degrees of Freedom, Full Model})} = 4.781$$



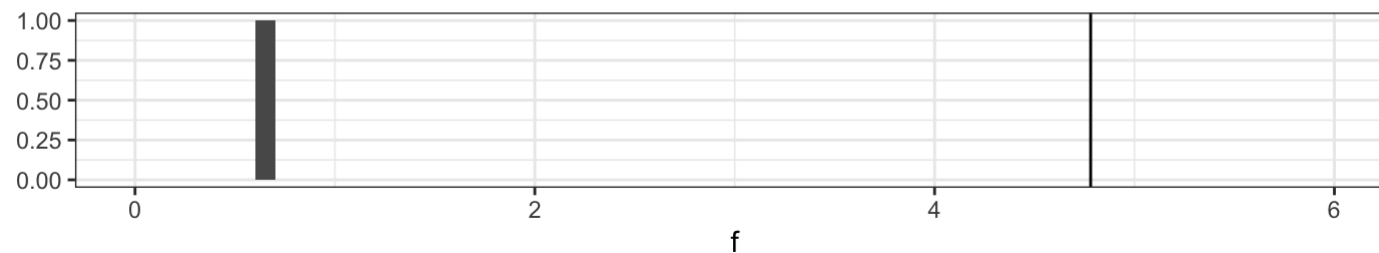
Heading towards a p-value

p-value: If H_0 is correct, what proportion of samples would give you F statistics at least as extreme as the F statistic of 4.78 we got from our actual data?

Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



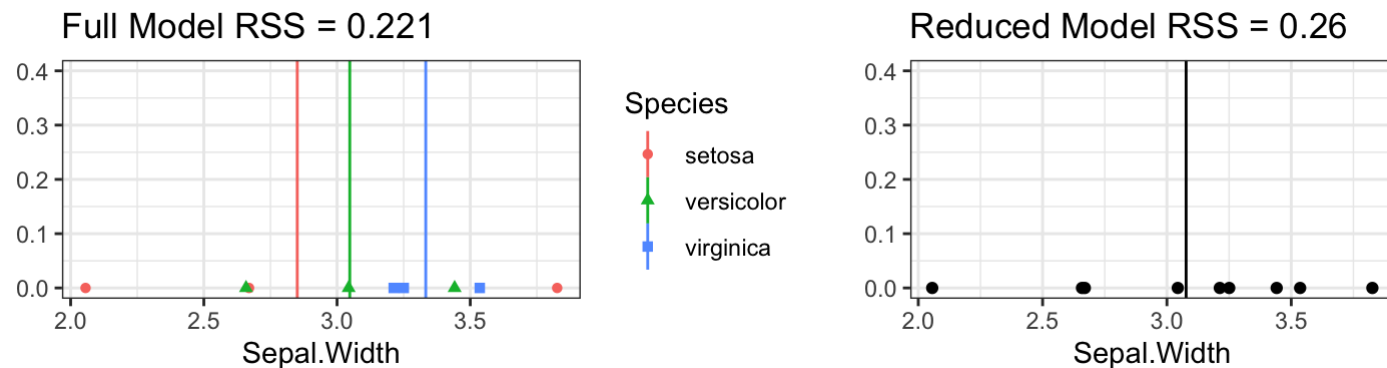
$$F = \frac{(\text{Extra Sum of Squares}) / (\text{Extra Degrees of Freedom})}{(\text{Residual Sum of Squares, Full Model}) / (\text{Degrees of Freedom, Full Model})} = 0.695$$



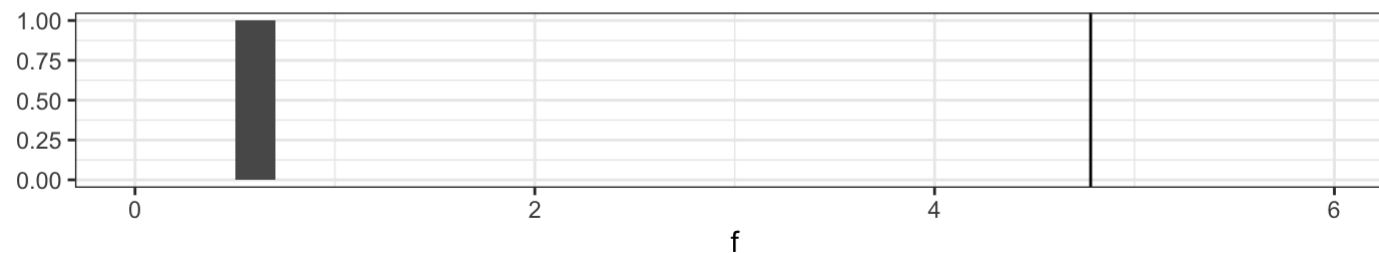
Heading towards a p-value

p-value: If H_0 is correct, what proportion of samples would give you F statistics at least as extreme as the F statistic of 4.78 we got from our actual data?

Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



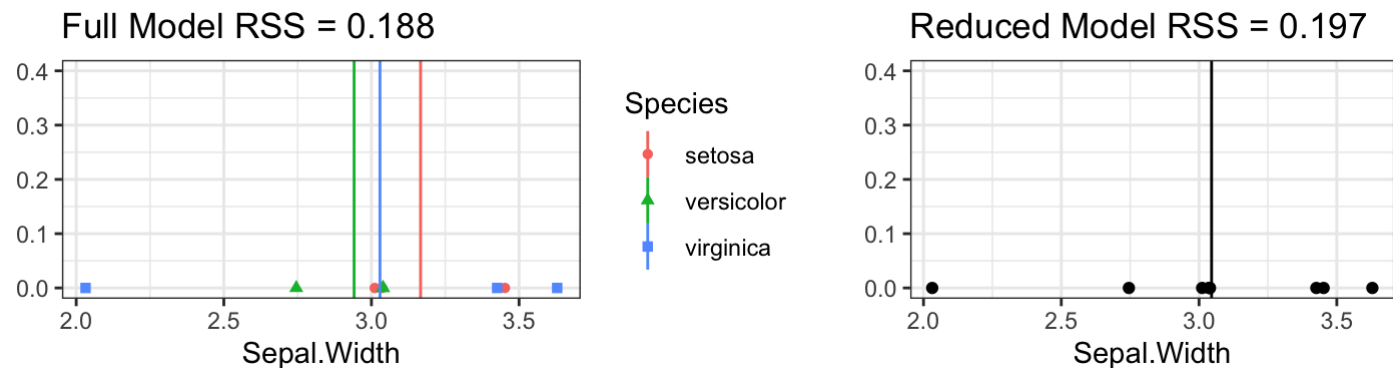
$$F = \frac{(\text{Extra Sum of Squares})/(\text{Extra Degrees of Freedom})}{(\text{Residual Sum of Squares, Full Model})/(\text{Degrees of Freedom, Full Model})} = 0.533$$



Heading towards a p-value

p-value: If H_0 is correct, what proportion of samples would give you F statistics at least as extreme as the F statistic of 4.78 we got from our actual data?

Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



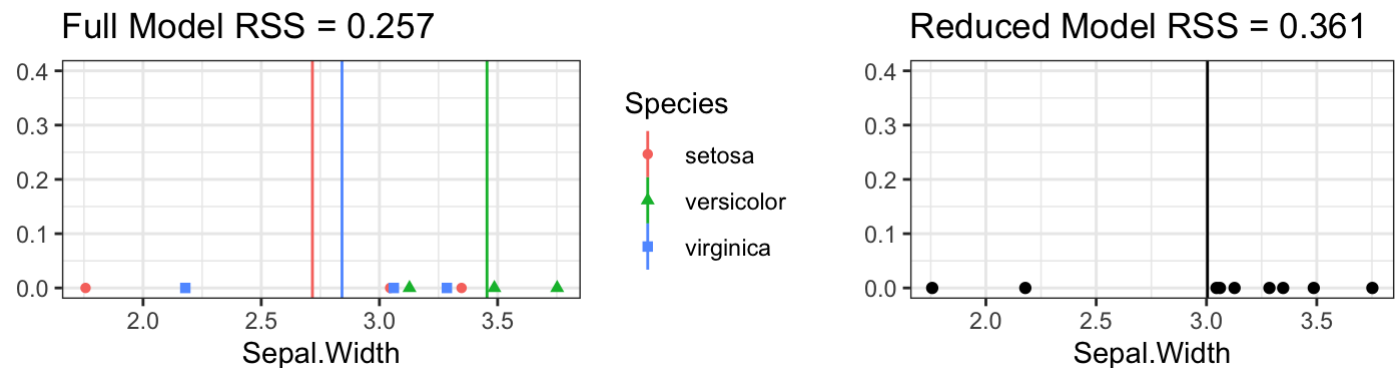
$$F = \frac{(\text{Extra Sum of Squares})/(\text{Extra Degrees of Freedom})}{(\text{Residual Sum of Squares, Full Model})/(\text{Degrees of Freedom, Full Model})} = 0.137$$



Heading towards a p-value

p-value: If H_0 is correct, what proportion of samples would give you F statistics at least as extreme as the F statistic of 4.78 we got from our actual data?

Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



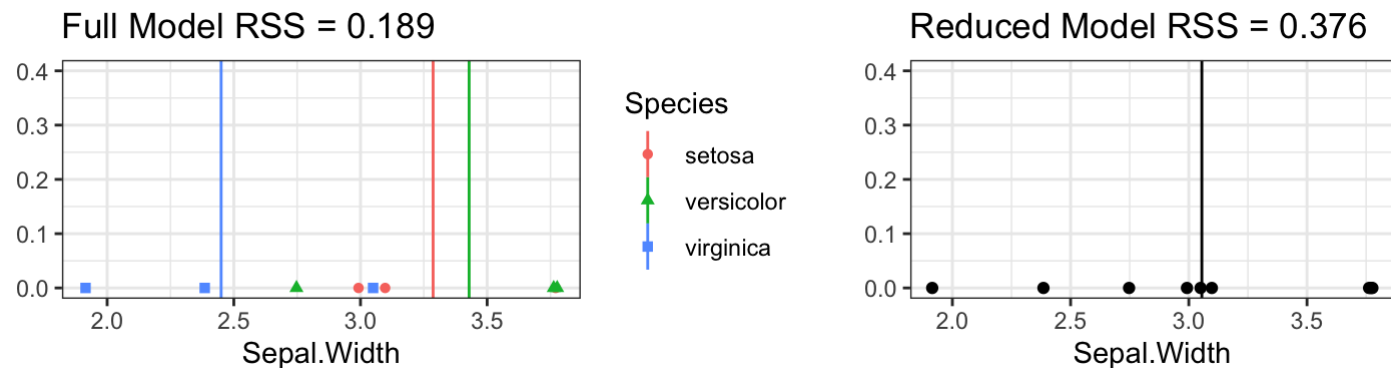
$$F = \frac{(\text{Extra Sum of Squares})/(\text{Extra Degrees of Freedom})}{(\text{Residual Sum of Squares, Full Model})/(\text{Degrees of Freedom, Full Model})} = 1.218$$



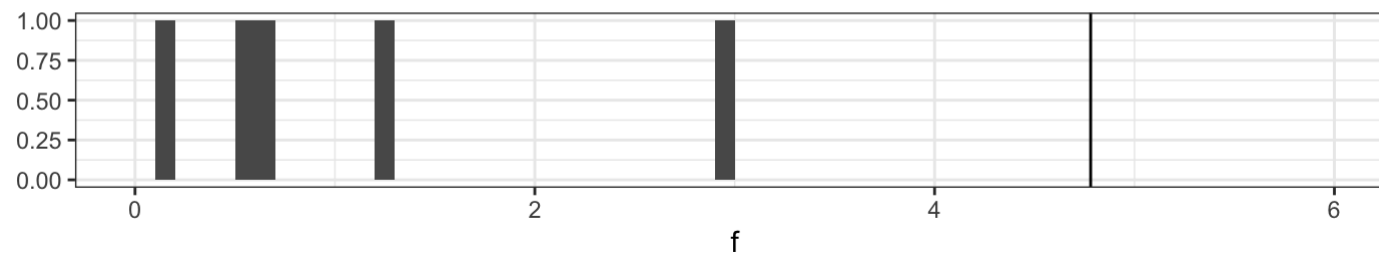
Heading towards a p-value

p-value: If H_0 is correct, what proportion of samples would give you F statistics at least as extreme as the F statistic of 4.78 we got from our actual data?

Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



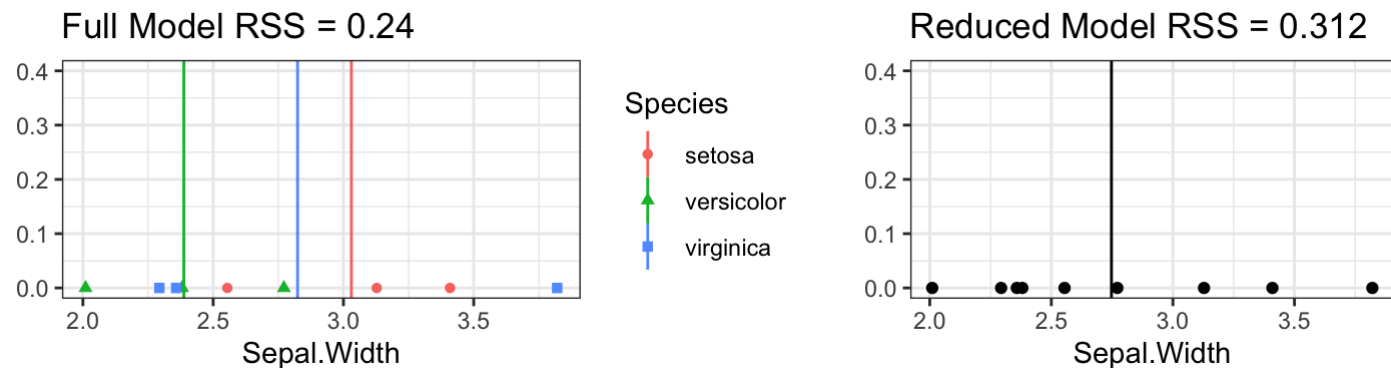
$$F = \frac{(\text{Extra Sum of Squares})/(\text{Extra Degrees of Freedom})}{(\text{Residual Sum of Squares, Full Model})/(\text{Degrees of Freedom, Full Model})} = 2.956$$



Heading towards a p-value

p-value: If H_0 is correct, what proportion of samples would give you F statistics at least as extreme as the F statistic of 4.78 we got from our actual data?

Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



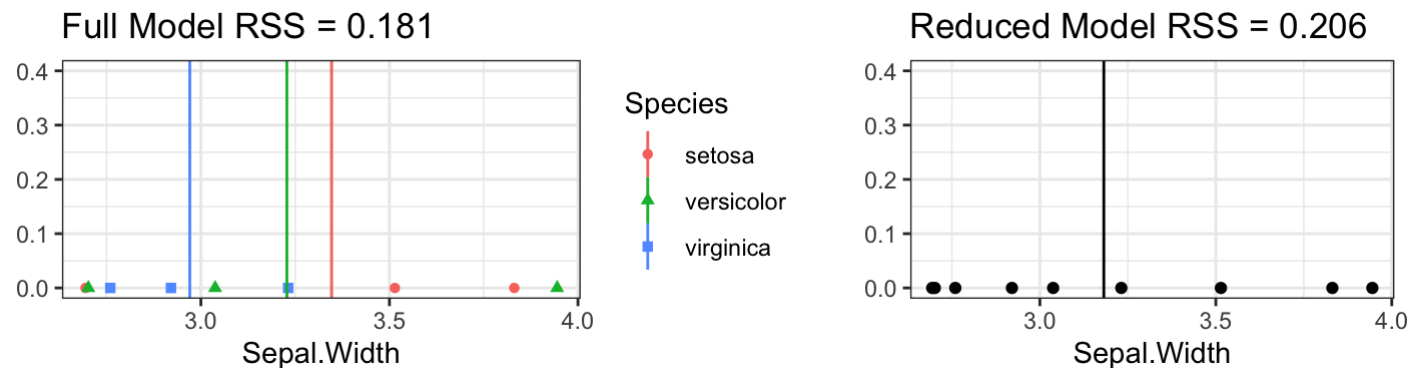
$$F = \frac{(\text{Extra Sum of Squares})/(\text{Extra Degrees of Freedom})}{(\text{Residual Sum of Squares, Full Model})/(\text{Degrees of Freedom, Full Model})} = 0.898$$



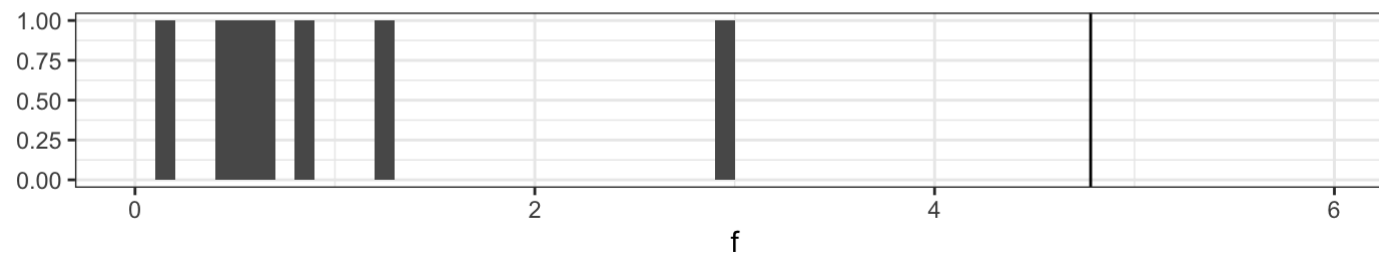
Heading towards a p-value

p-value: If H_0 is correct, what proportion of samples would give you F statistics at least as extreme as the F statistic of 4.78 we got from our actual data?

Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



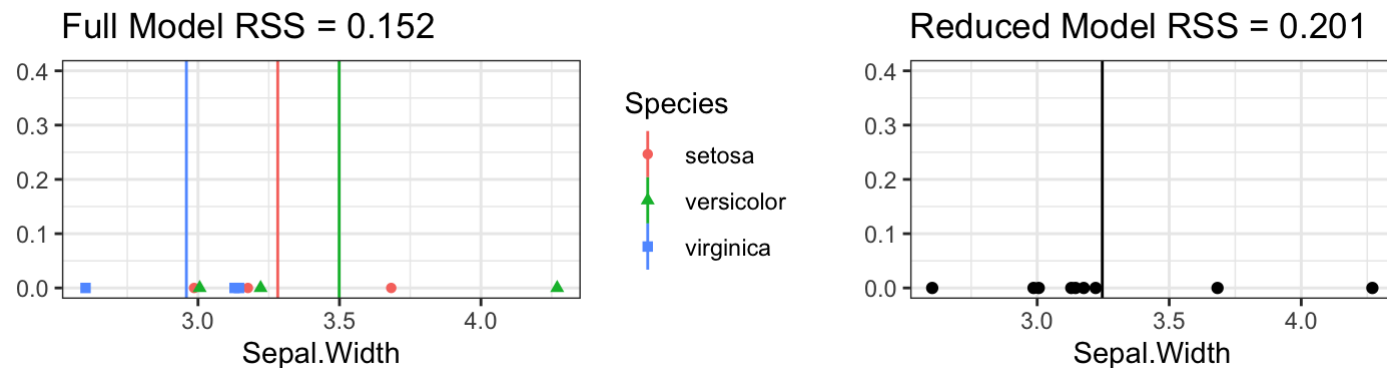
$$F = \frac{(\text{Extra Sum of Squares})/(\text{Extra Degrees of Freedom})}{(\text{Residual Sum of Squares, Full Model})/(\text{Degrees of Freedom, Full Model})} = 0.408$$



Heading towards a p-value

p-value: If H_0 is correct, what proportion of samples would give you F statistics at least as extreme as the F statistic of 4.78 we got from our actual data?

Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



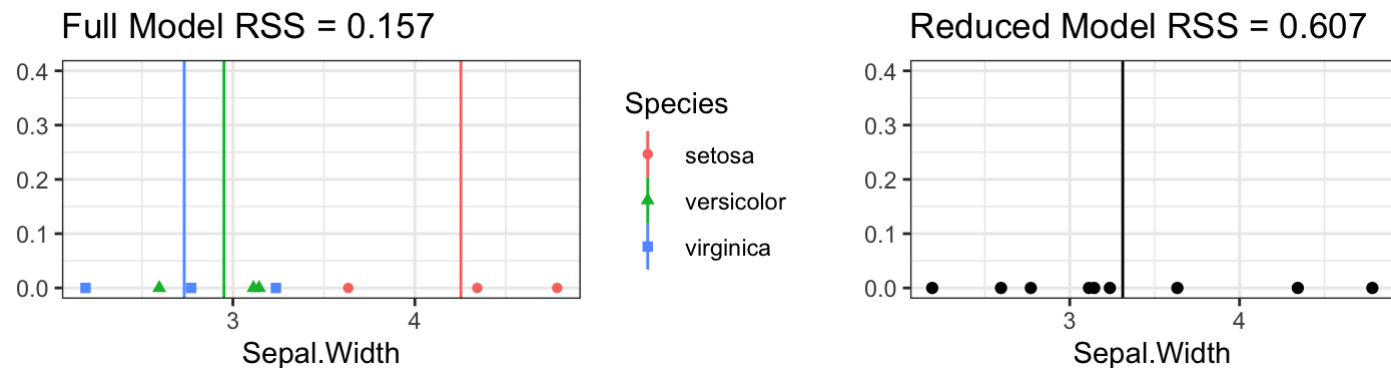
$$F = \frac{(\text{Extra Sum of Squares})/(\text{Extra Degrees of Freedom})}{(\text{Residual Sum of Squares, Full Model})/(\text{Degrees of Freedom, Full Model})} = 0.974$$



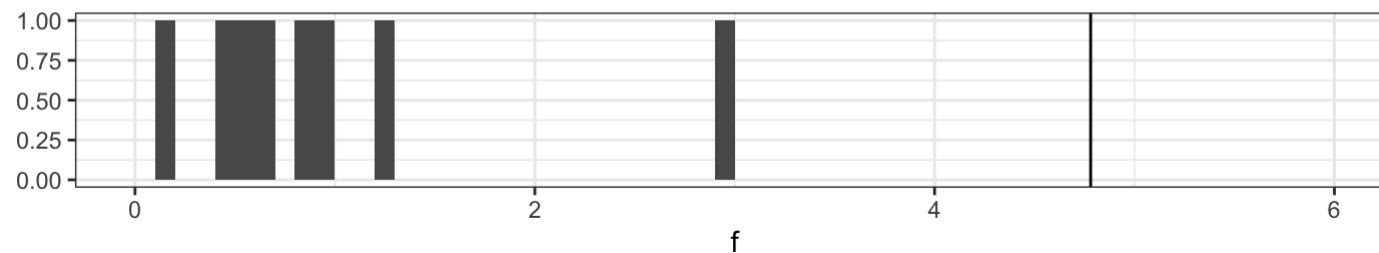
Heading towards a p-value

p-value: If H_0 is correct, what proportion of samples would give you F statistics at least as extreme as the F statistic of 4.78 we got from our actual data?

Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



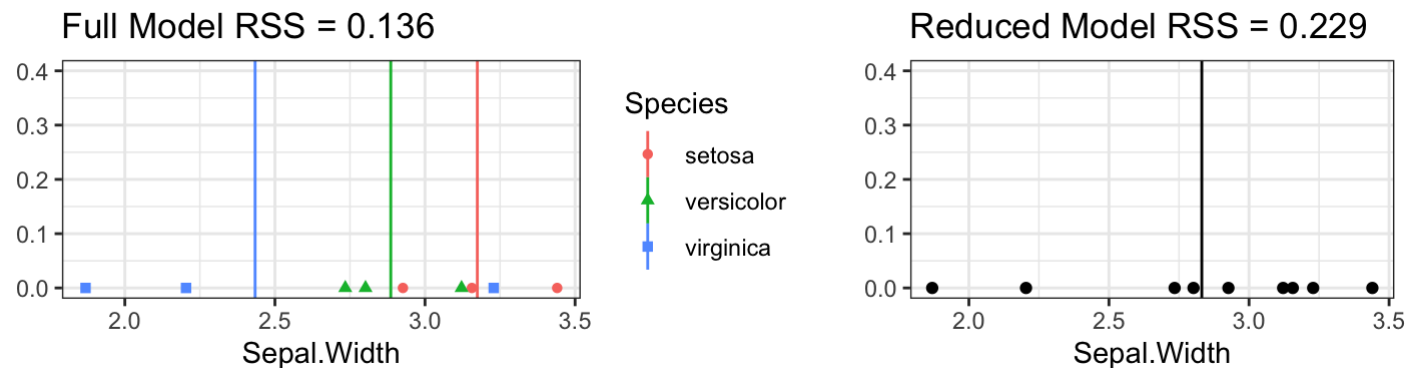
$$F = \frac{(\text{Extra Sum of Squares})/(\text{Extra Degrees of Freedom})}{(\text{Residual Sum of Squares, Full Model})/(\text{Degrees of Freedom, Full Model})} = 8.621$$



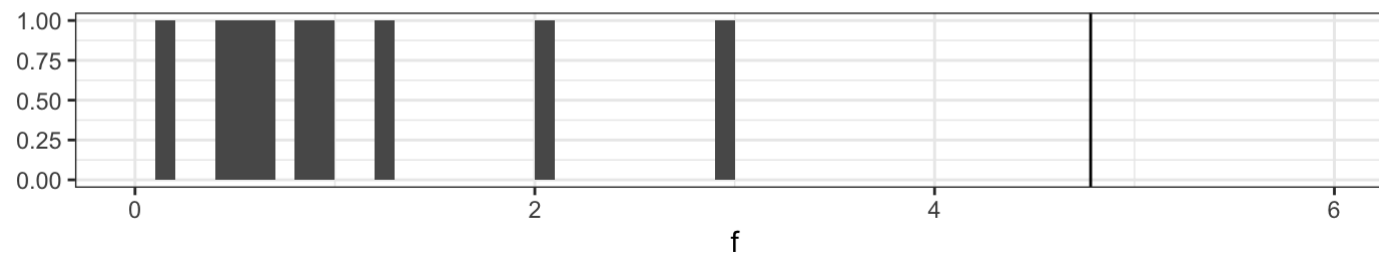
Heading towards a p-value

p-value: If H_0 is correct, what proportion of samples would give you F statistics at least as extreme as the F statistic of 4.78 we got from our actual data?

Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



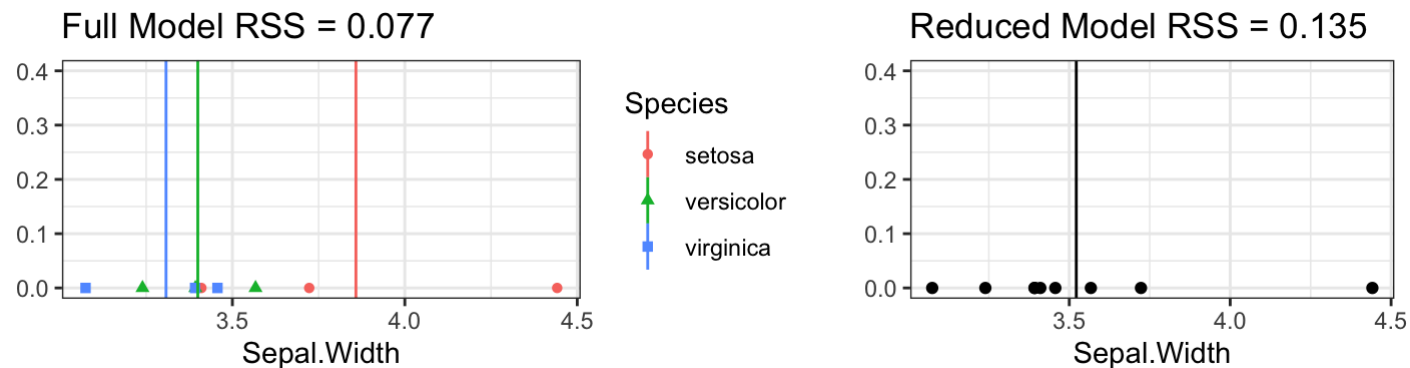
$$F = \frac{(\text{Extra Sum of Squares})/(\text{Extra Degrees of Freedom})}{(\text{Residual Sum of Squares, Full Model})/(\text{Degrees of Freedom, Full Model})} = 2.051$$



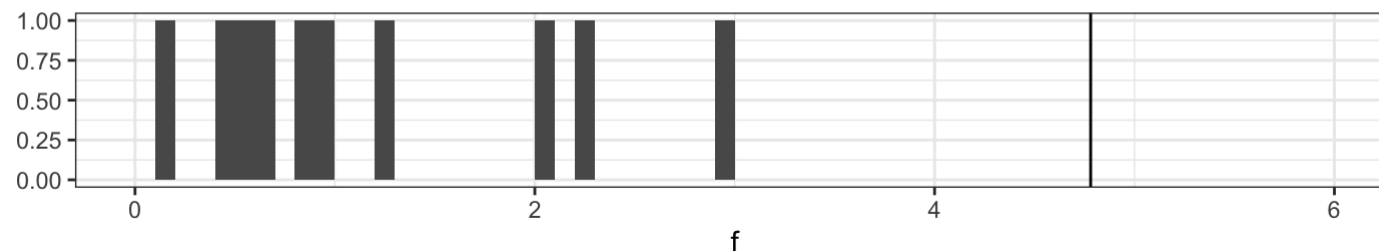
Heading towards a p-value

p-value: If H_0 is correct, what proportion of samples would give you F statistics at least as extreme as the F statistic of 4.78 we got from our actual data?

Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



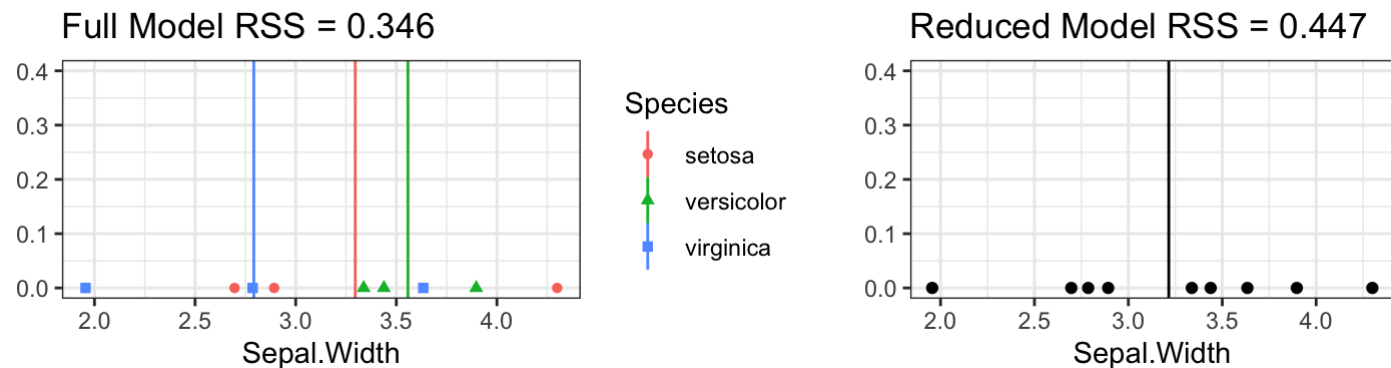
$$F = \frac{(\text{Extra Sum of Squares})/(\text{Extra Degrees of Freedom})}{(\text{Residual Sum of Squares, Full Model})/(\text{Degrees of Freedom, Full Model})} = 2.246$$



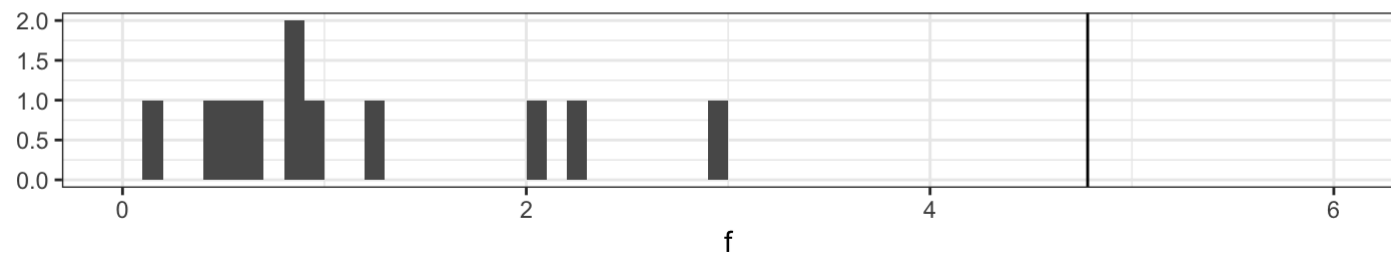
Heading towards a p-value

p-value: If H_0 is correct, what proportion of samples would give you F statistics at least as extreme as the F statistic of 4.78 we got from our actual data?

Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



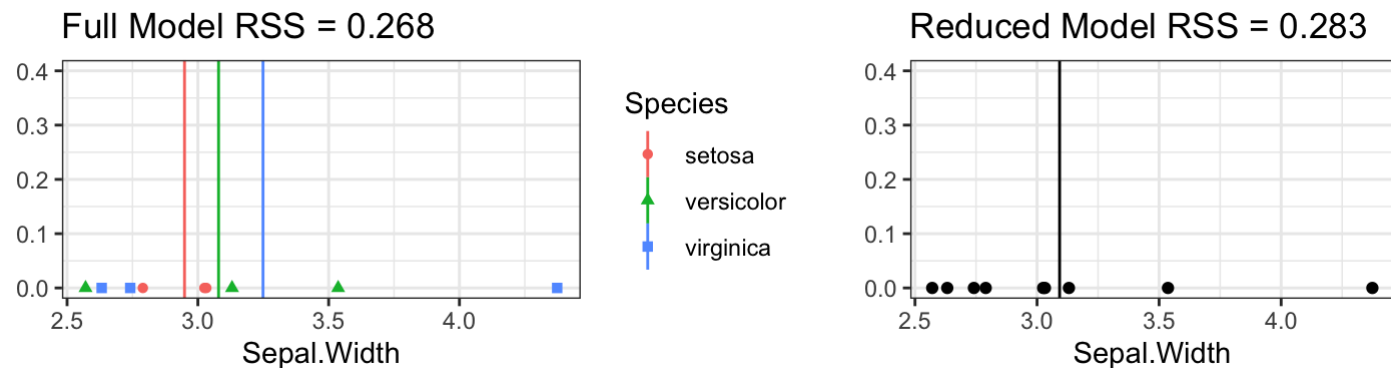
$$F = \frac{(\text{Extra Sum of Squares})/(\text{Extra Degrees of Freedom})}{(\text{Residual Sum of Squares, Full Model})/(\text{Degrees of Freedom, Full Model})} = 0.876$$



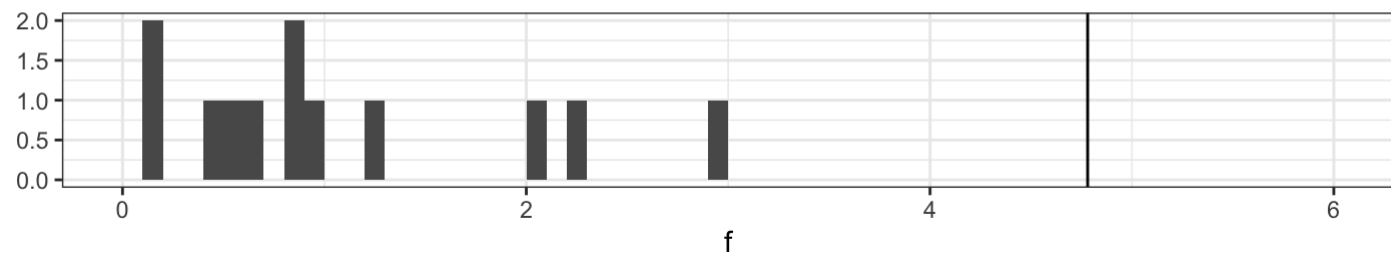
Heading towards a p-value

p-value: If H_0 is correct, what proportion of samples would give you F statistics at least as extreme as the F statistic of 4.78 we got from our actual data?

Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



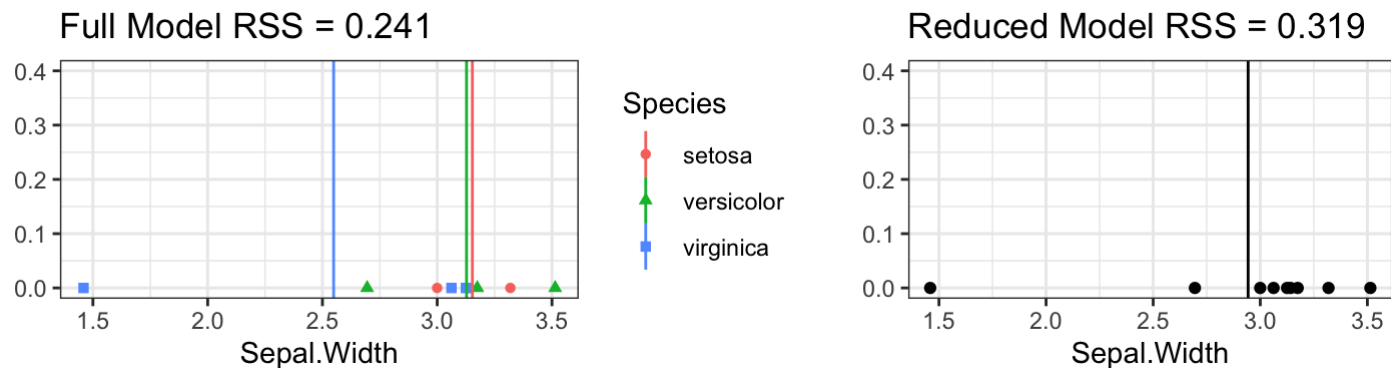
$$F = \frac{(\text{Extra Sum of Squares})/(\text{Extra Degrees of Freedom})}{(\text{Residual Sum of Squares, Full Model})/(\text{Degrees of Freedom, Full Model})} = 0.169$$



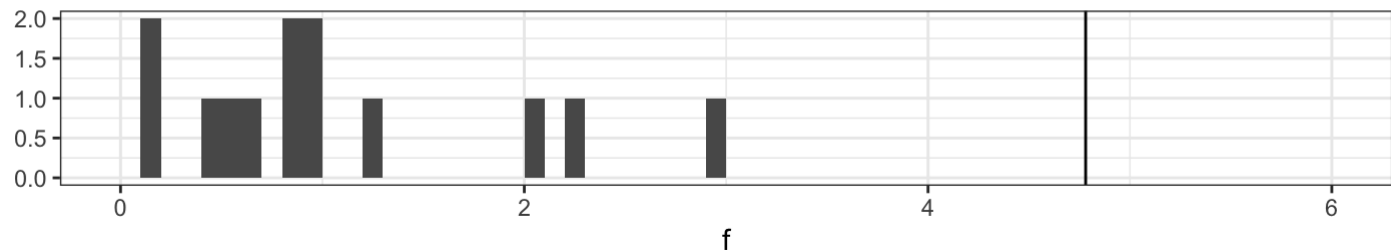
Heading towards a p-value

p-value: If H_0 is correct, what proportion of samples would give you F statistics at least as extreme as the F statistic of 4.78 we got from our actual data?

Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



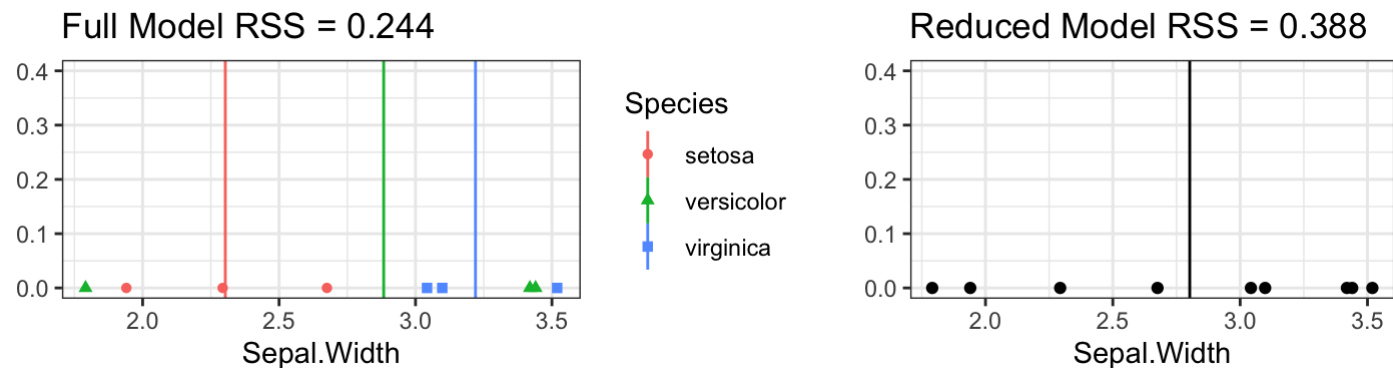
$$F = \frac{(\text{Extra Sum of Squares})/(\text{Extra Degrees of Freedom})}{(\text{Residual Sum of Squares, Full Model})/(\text{Degrees of Freedom, Full Model})} = 0.968$$



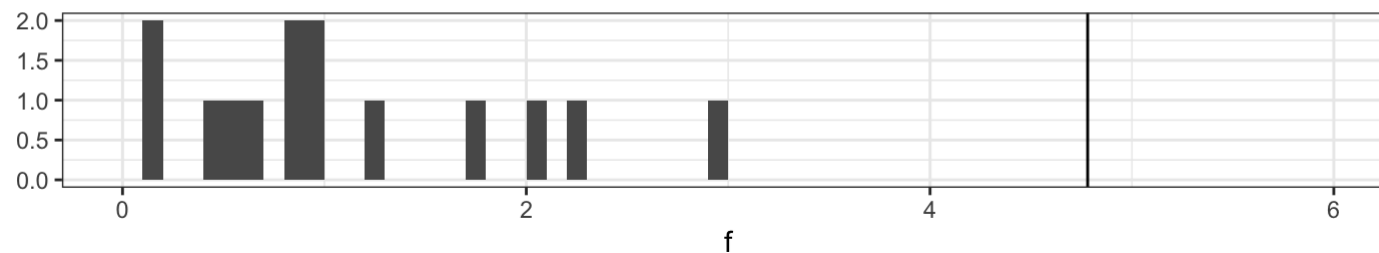
Heading towards a p-value

p-value: If H_0 is correct, what proportion of samples would give you F statistics at least as extreme as the F statistic of 4.78 we got from our actual data?

Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



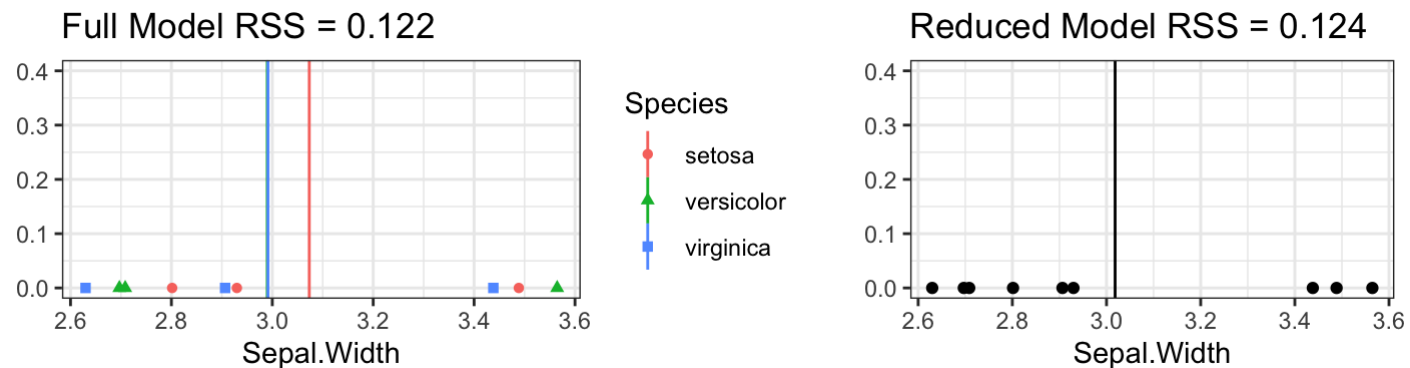
$$F = \frac{(\text{Extra Sum of Squares}) / (\text{Extra Degrees of Freedom})}{(\text{Residual Sum of Squares, Full Model}) / (\text{Degrees of Freedom, Full Model})} = 1.762$$



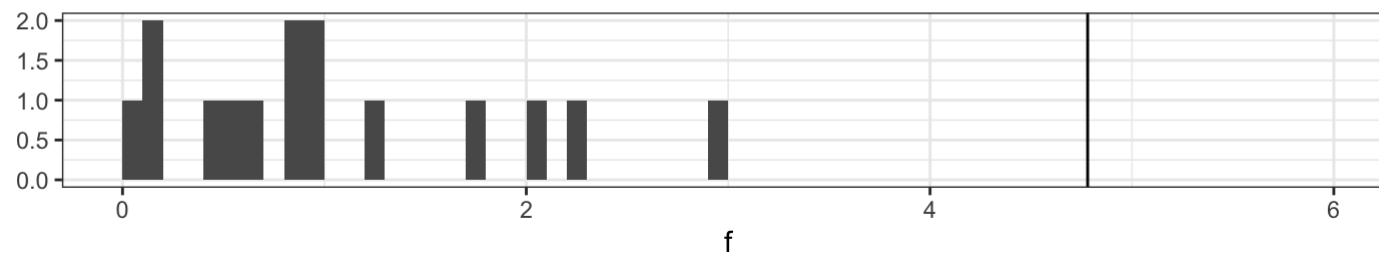
Heading towards a p-value

p-value: If H_0 is correct, what proportion of samples would give you F statistics at least as extreme as the F statistic of 4.78 we got from our actual data?

Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



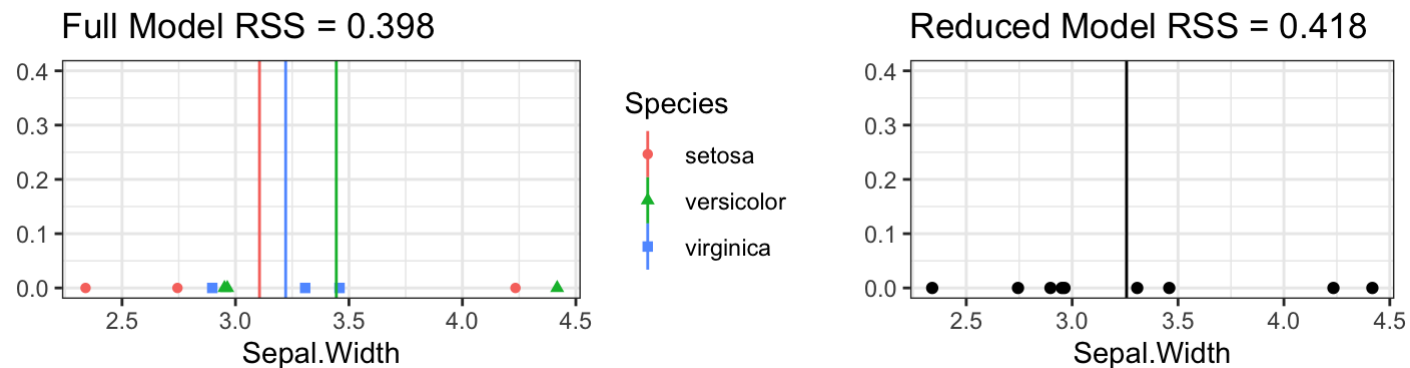
$$F = \frac{(\text{Extra Sum of Squares})/(\text{Extra Degrees of Freedom})}{(\text{Residual Sum of Squares, Full Model})/(\text{Degrees of Freedom, Full Model})} = 0.037$$



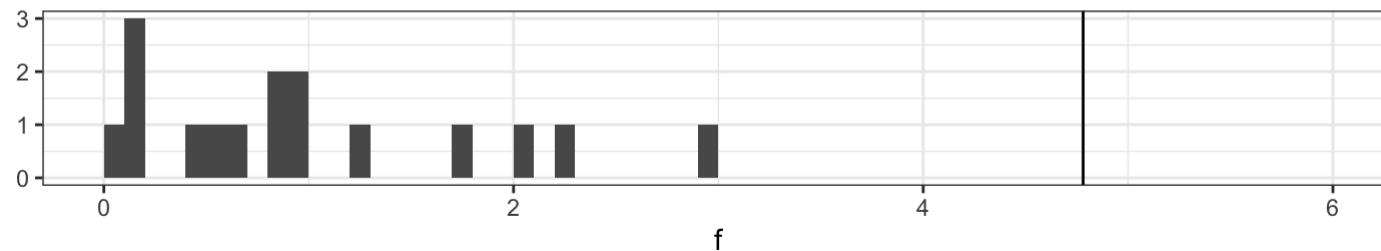
Heading towards a p-value

p-value: If H_0 is correct, what proportion of samples would give you F statistics at least as extreme as the F statistic of 4.78 we got from our actual data?

Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



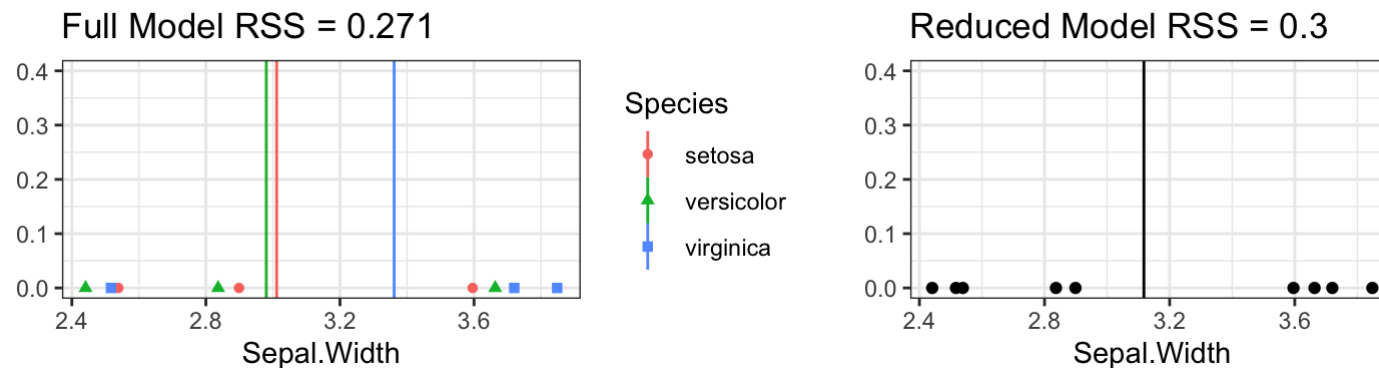
$$F = \frac{(\text{Extra Sum of Squares})/(\text{Extra Degrees of Freedom})}{(\text{Residual Sum of Squares, Full Model})/(\text{Degrees of Freedom, Full Model})} = 0.149$$



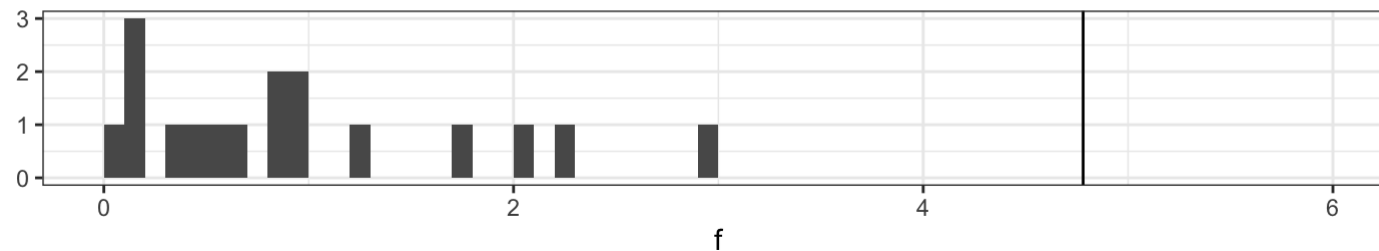
Heading towards a p-value

p-value: If H_0 is correct, what proportion of samples would give you F statistics at least as extreme as the F statistic of 4.78 we got from our actual data?

Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



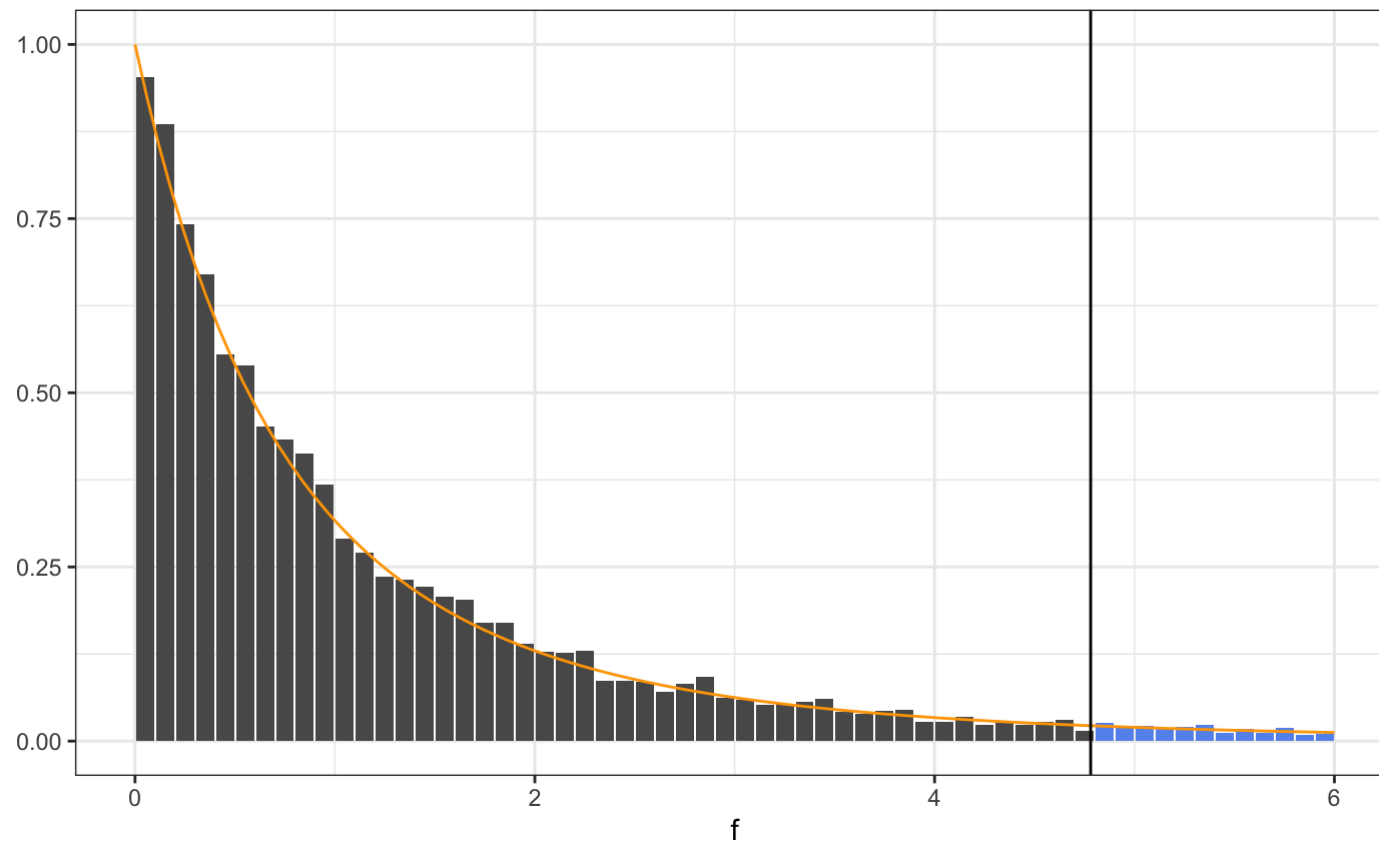
$$F = \frac{(\text{Extra Sum of Squares}) / (\text{Extra Degrees of Freedom})}{(\text{Residual Sum of Squares, Full Model}) / (\text{Degrees of Freedom, Full Model})} = 0.332$$



Heading towards a p-value

p-value: If H_0 is correct, what proportion of samples would give you F statistics at least as extreme as the F statistic of 4.78 we got from our actual data?

F statistics from 10,000 more samples, all generated assuming H_0 is true. The orange curve shows the theoretically derived F distribution.



A p-value

- The p-value is 0.05732
- If the null hypothesis that all three species had the same population mean were true, about 5.7% of samples would give you F statistics at least as extreme as the F statistic of 4.78 we got from our actual data. (Those samples contribute to the blue part of the histogram.)
- This offers some evidence against H_0 , but not very strong.

