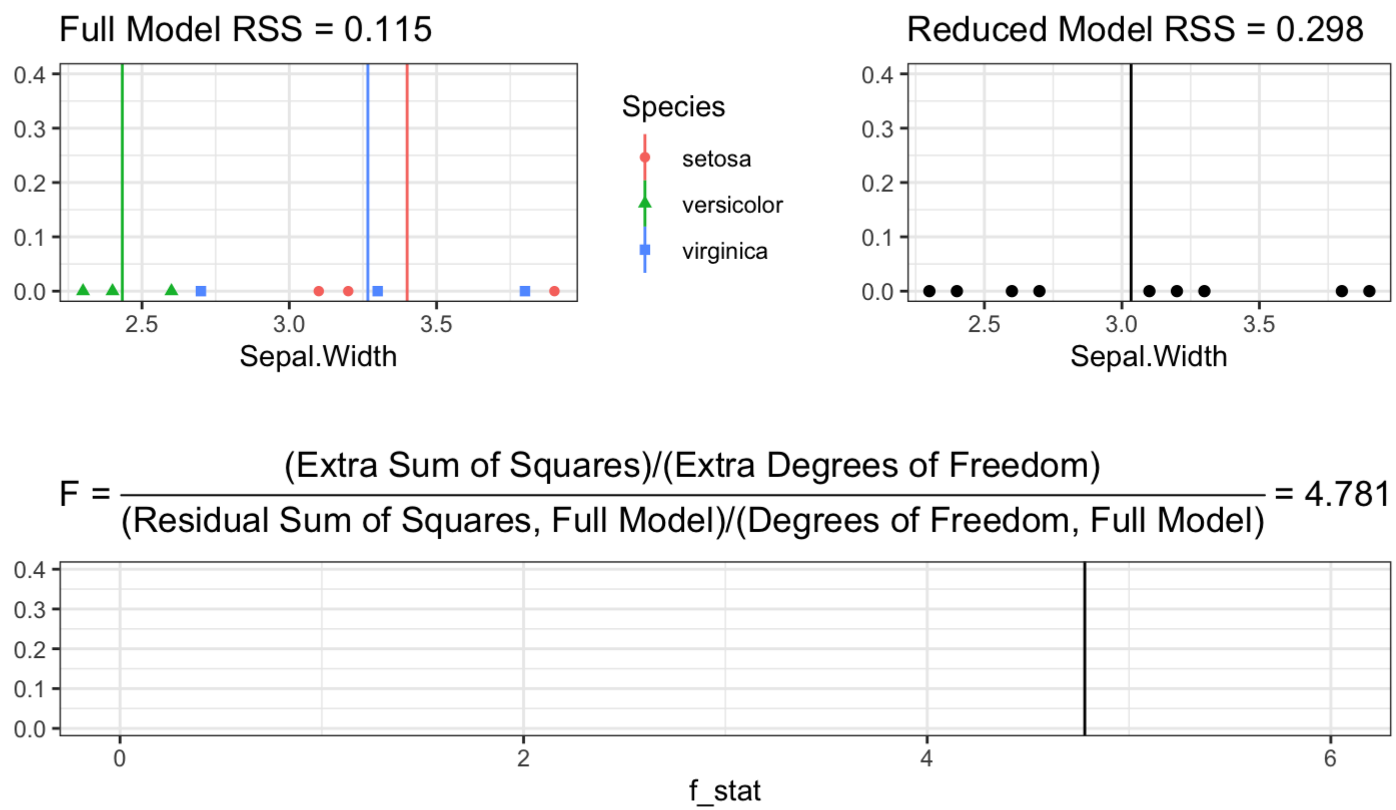
F Test p-values

Evan L. Ray

September 23, 2019

F tests: Iris example

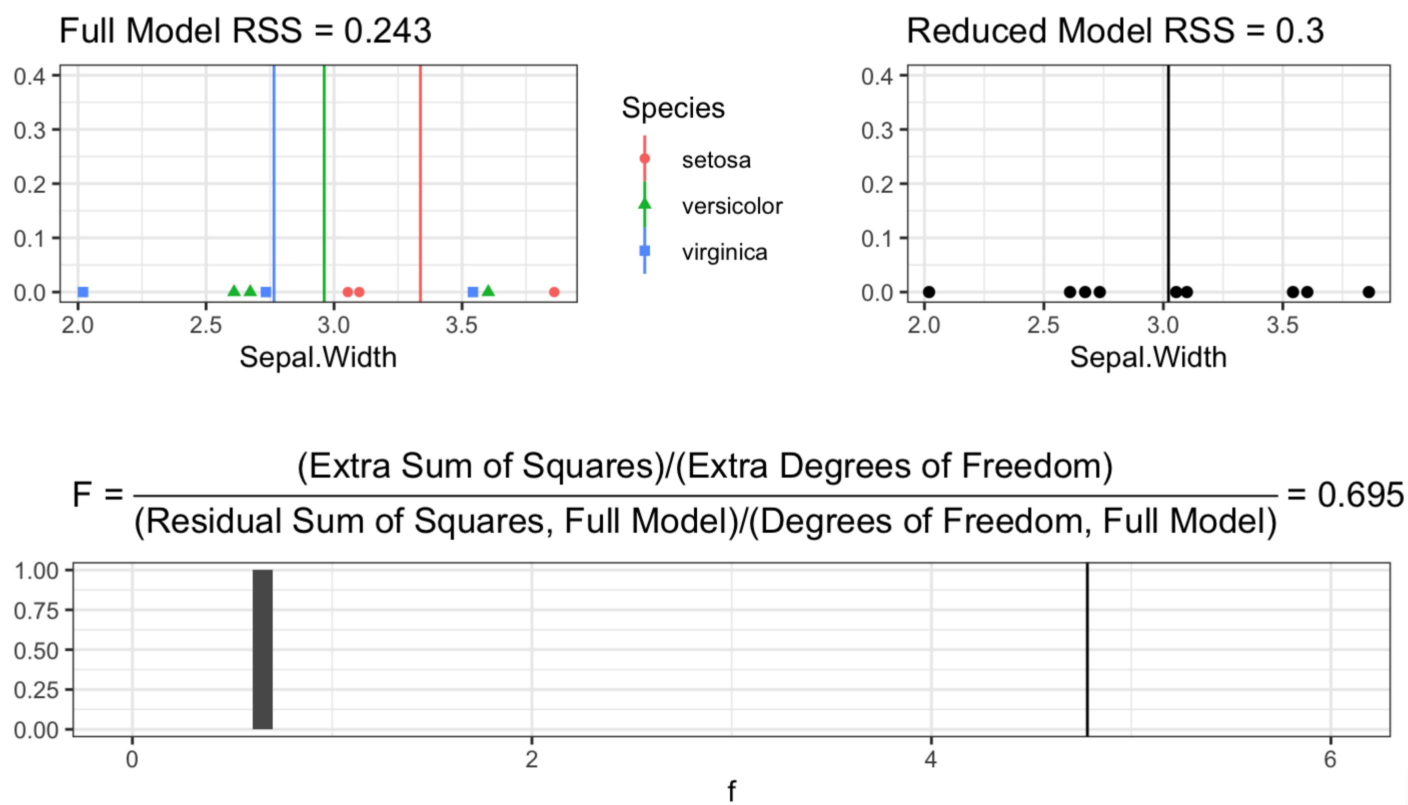
Suppose we have 3 flowers of each iris species, and we want to conduct an F test of $H_0 : \mu_1 = \mu_2 = \mu_3$.



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?

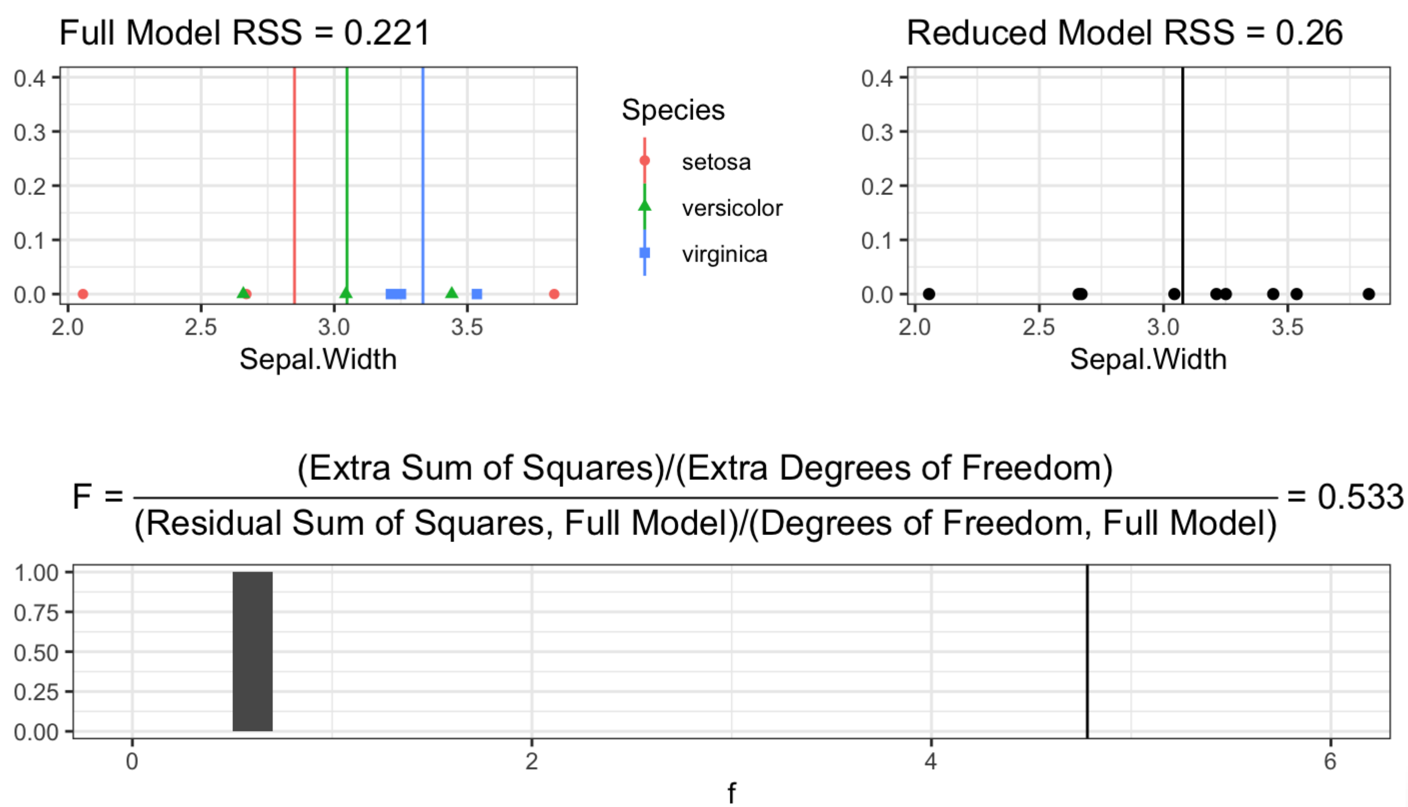
Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



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?

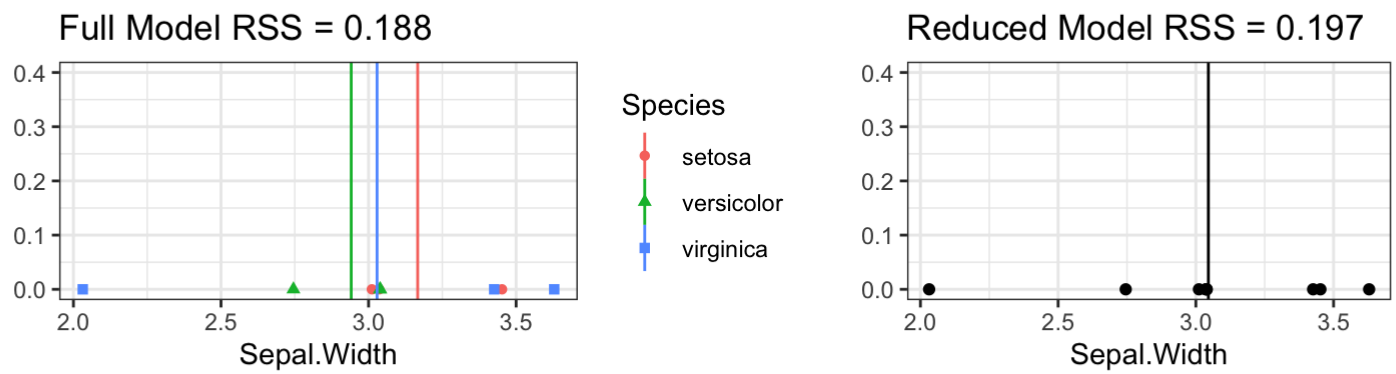
Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.

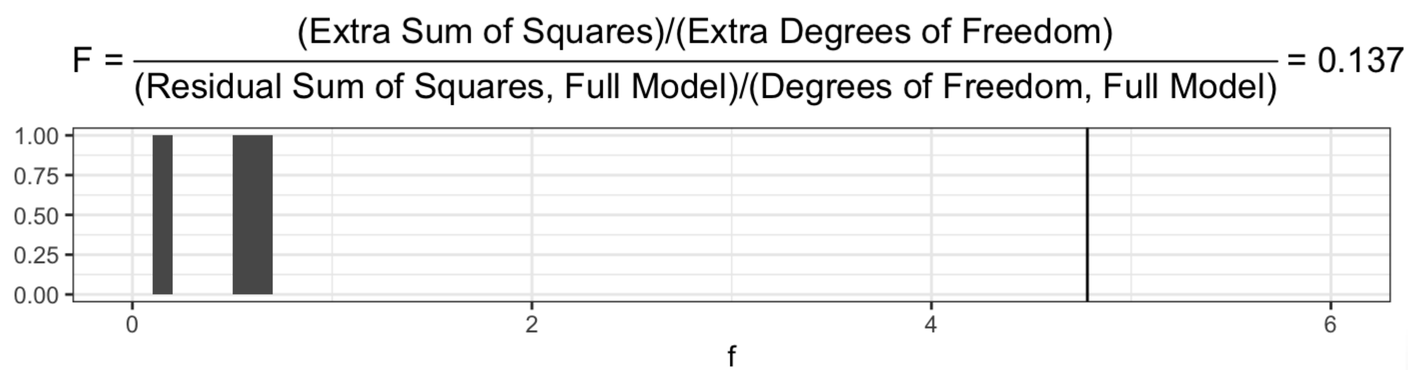


Heading towards a p-value

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.

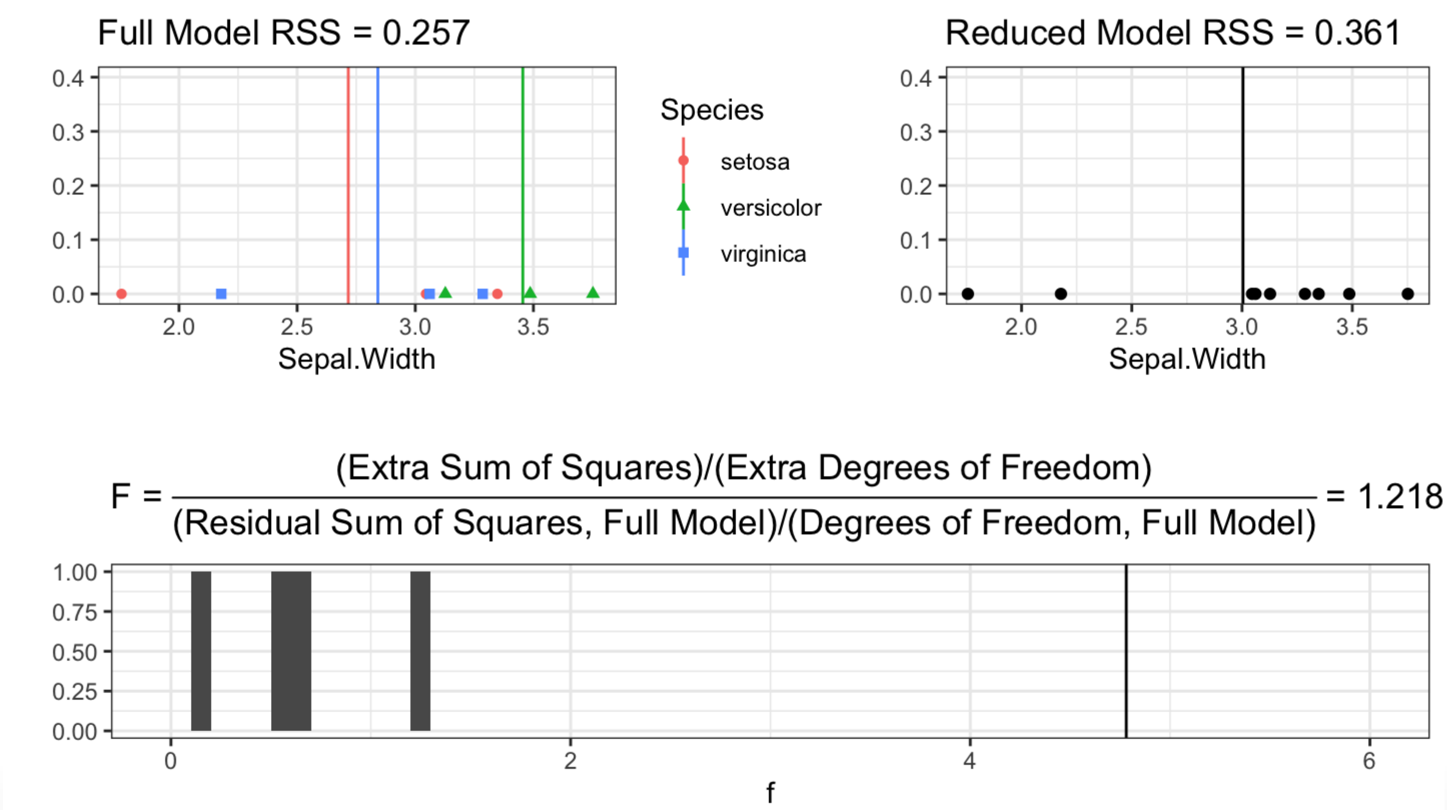




Heading towards a p-value

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?

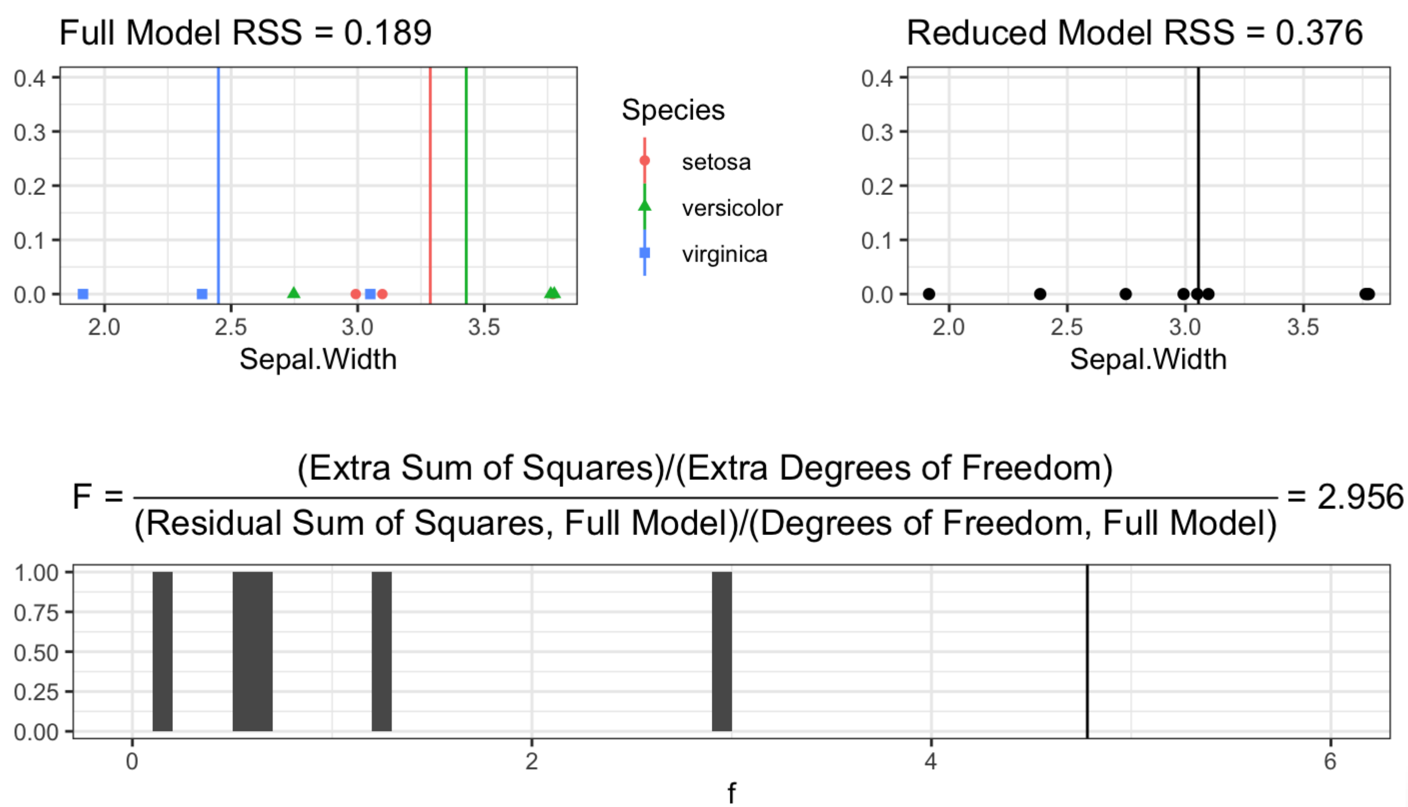
Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



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?

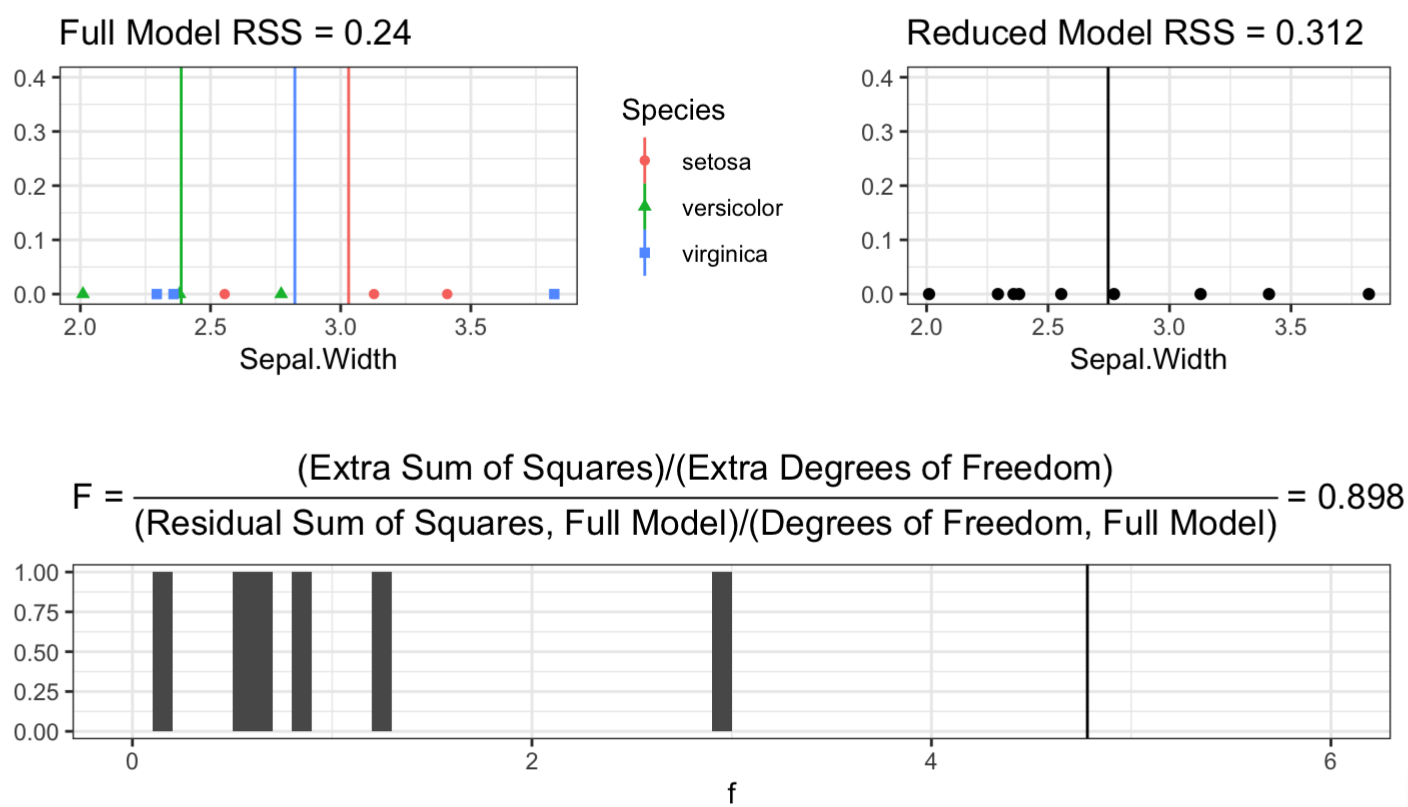
Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



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?

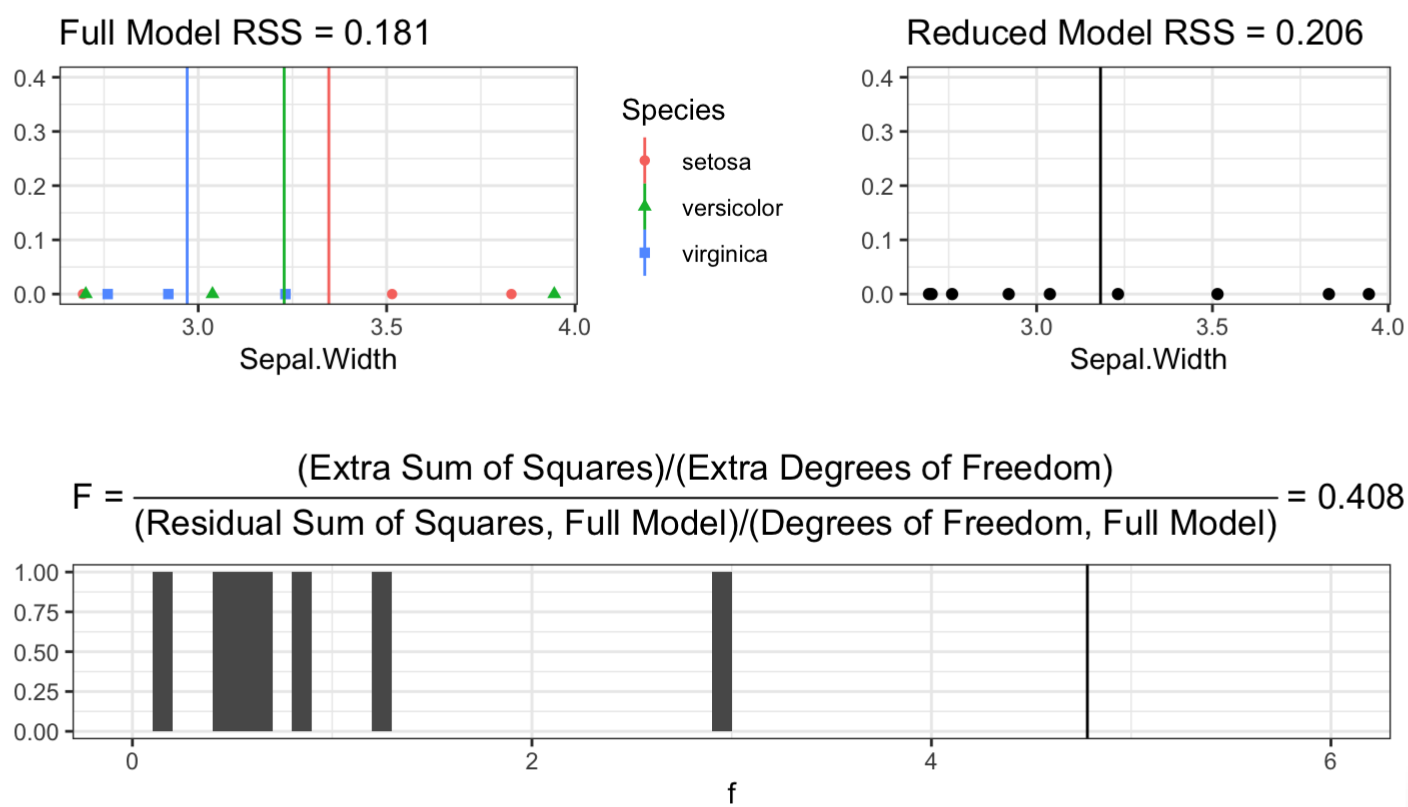
Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



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?

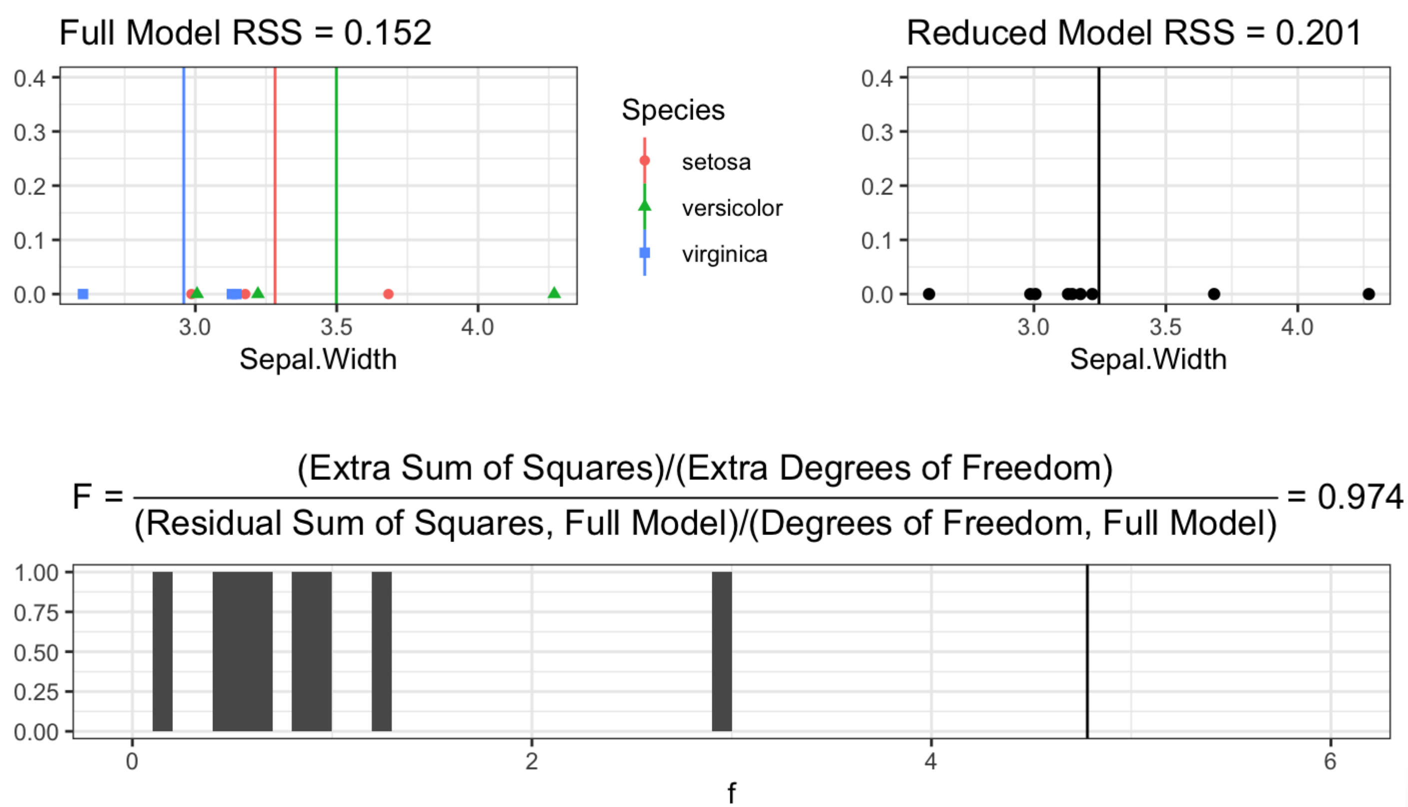
Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



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?

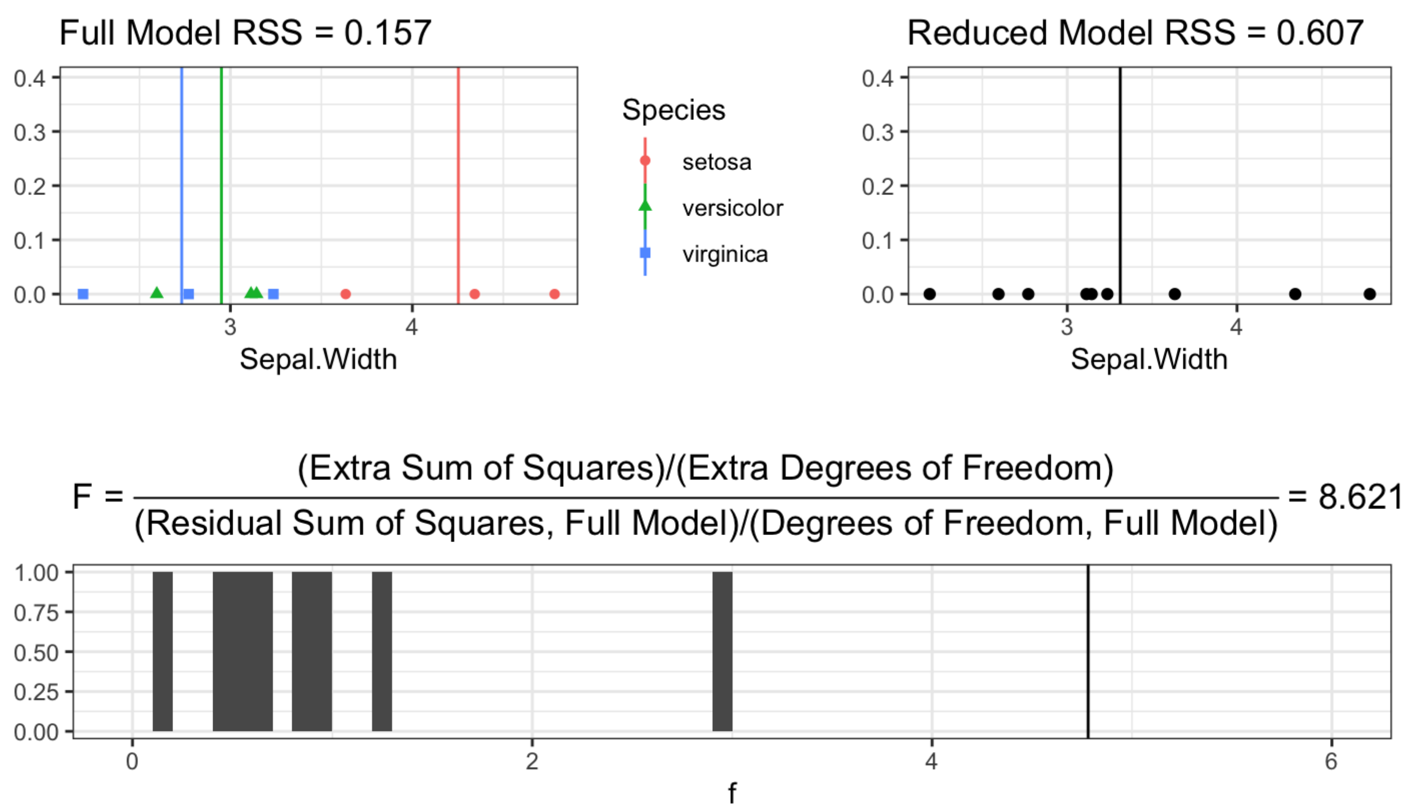
Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



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?

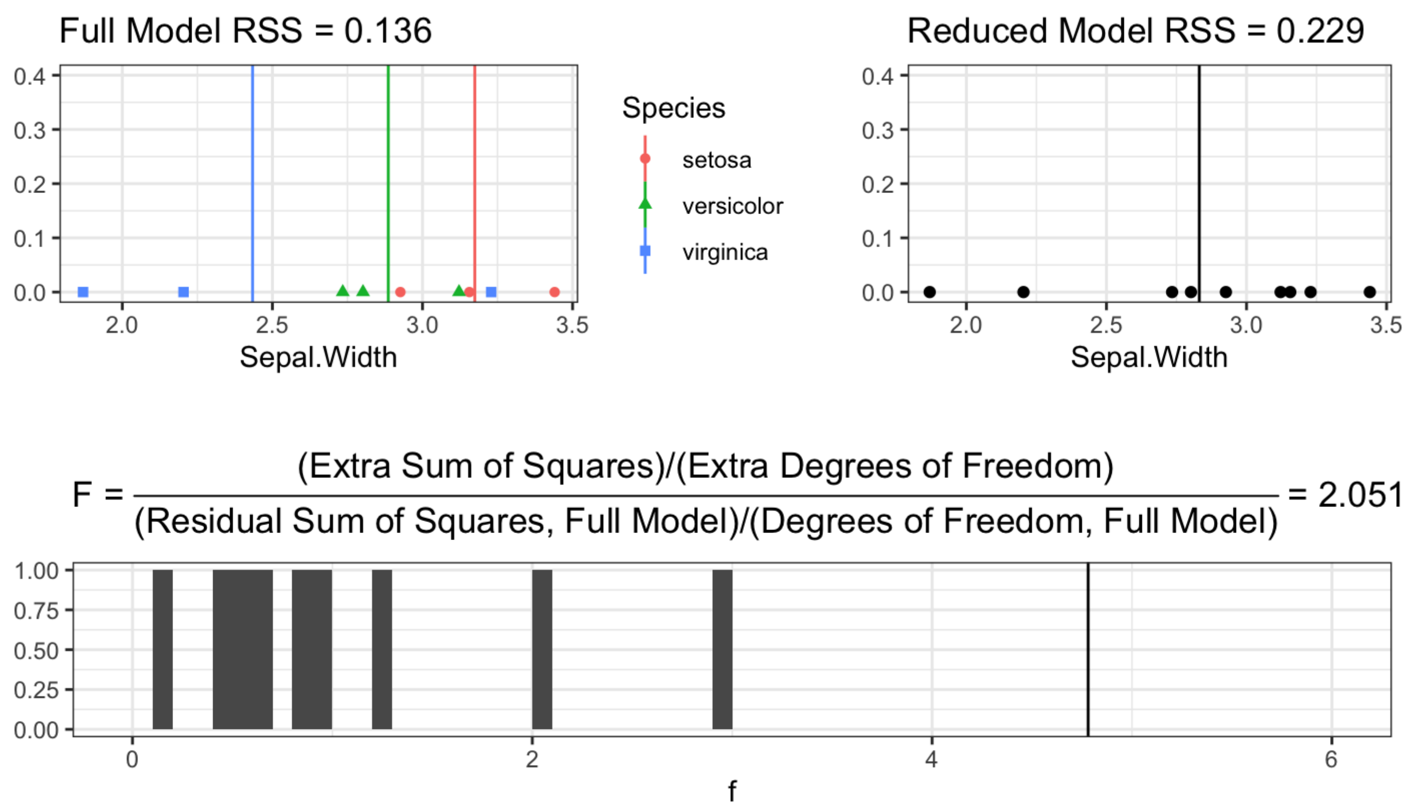
Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



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?

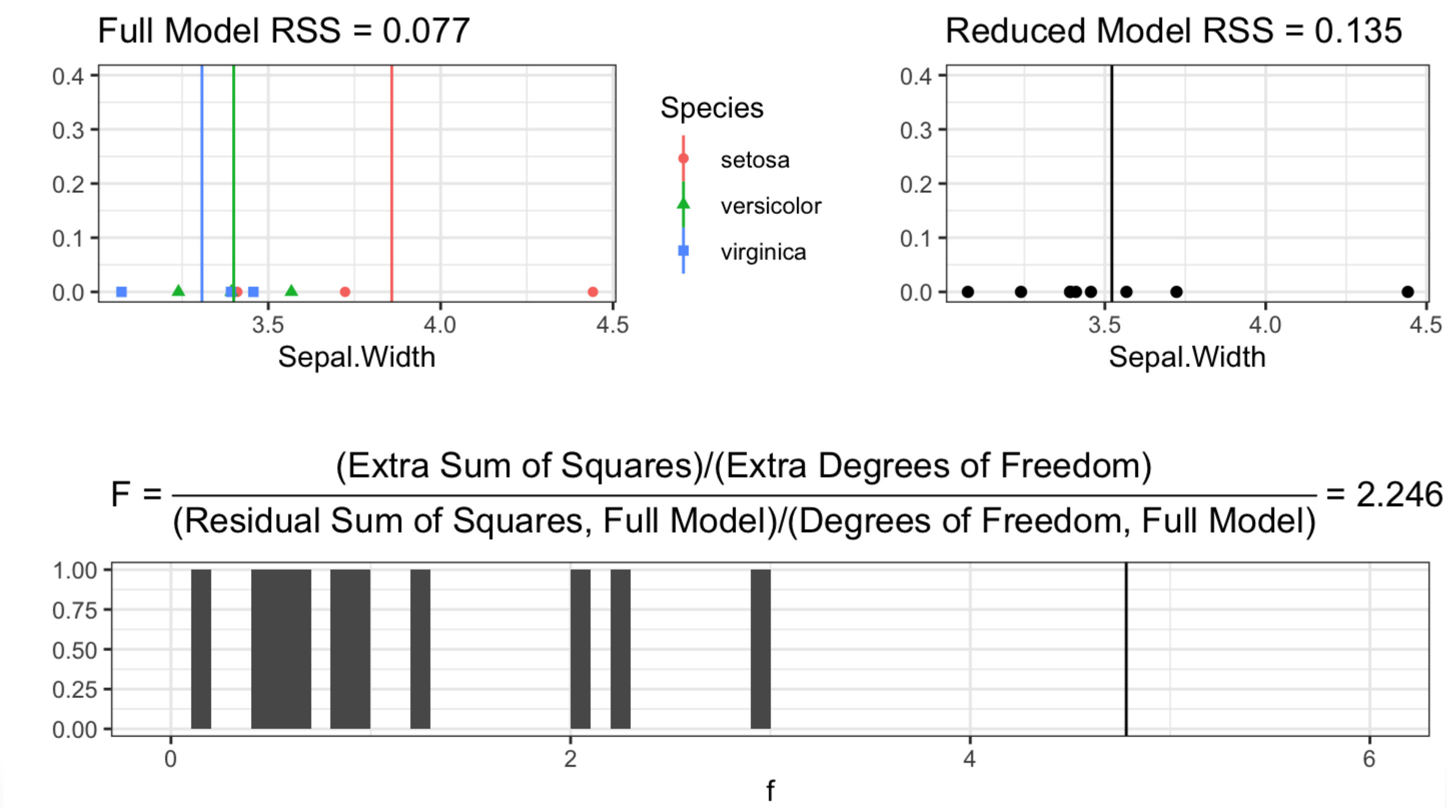
Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



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?

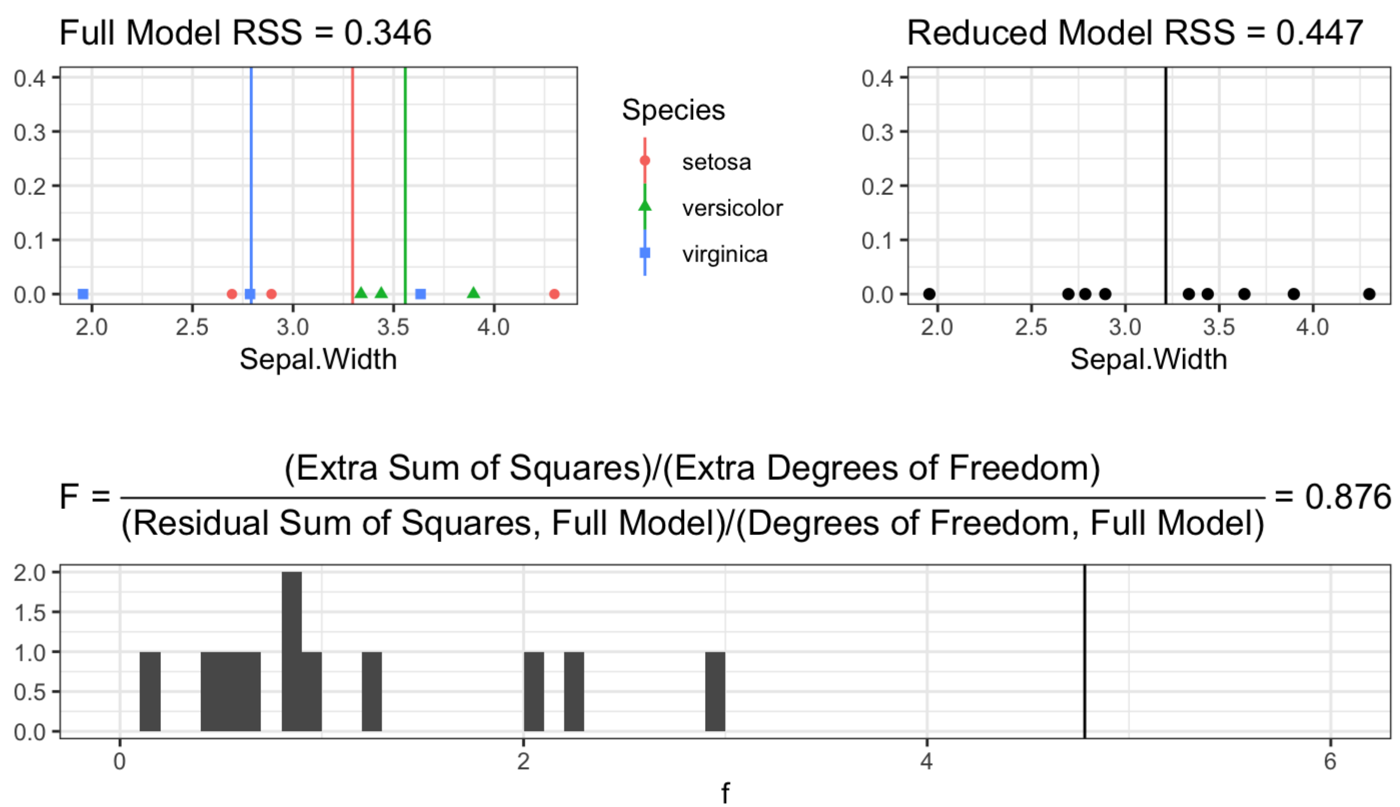
Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



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?

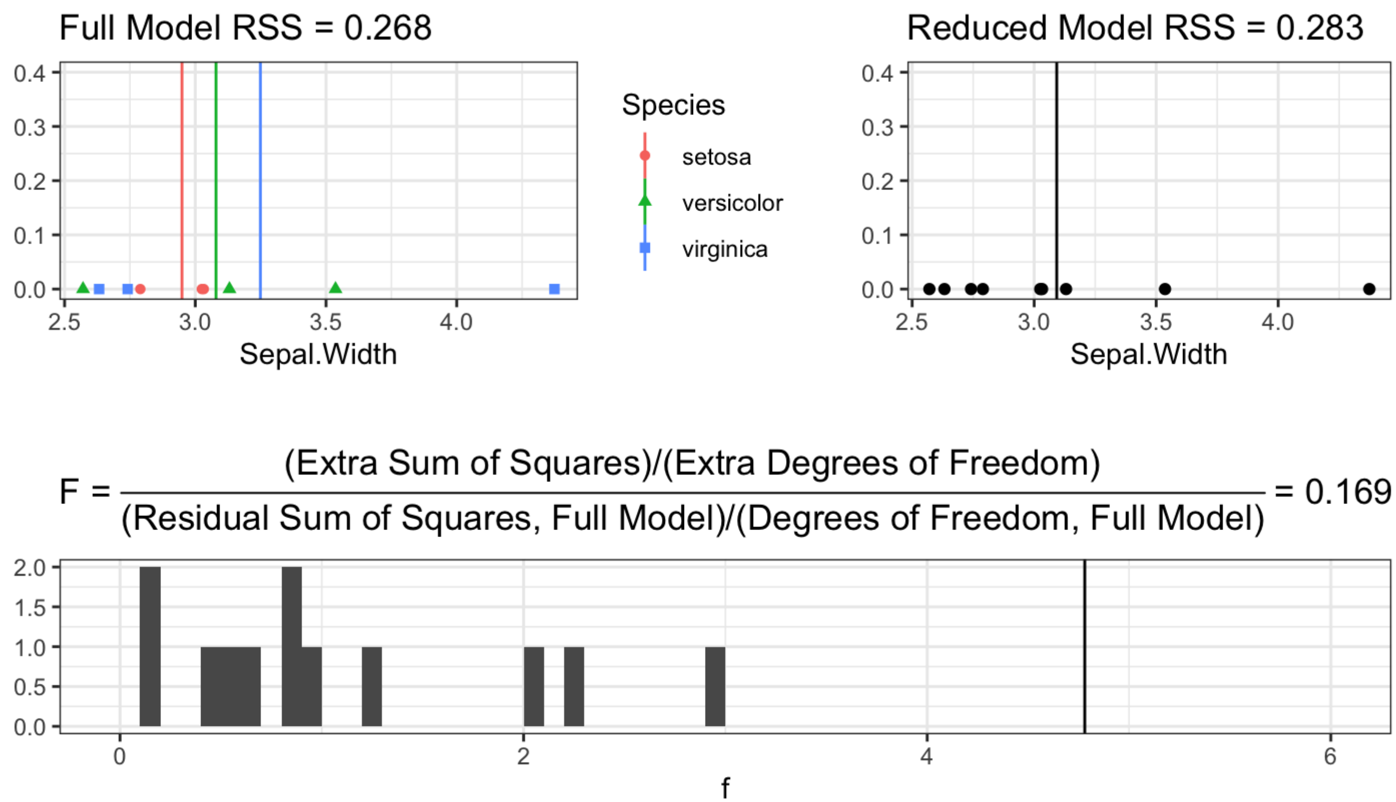
Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



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?

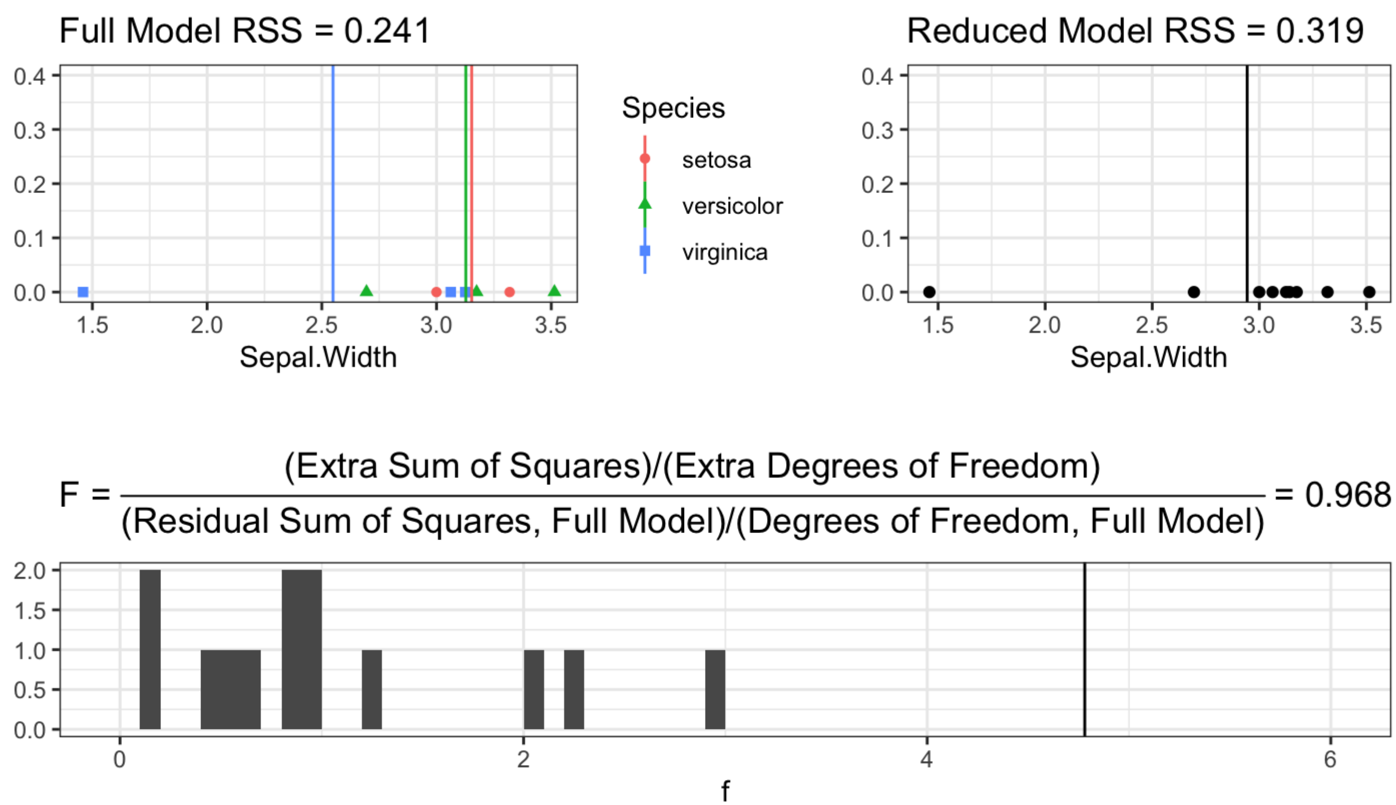
Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



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?

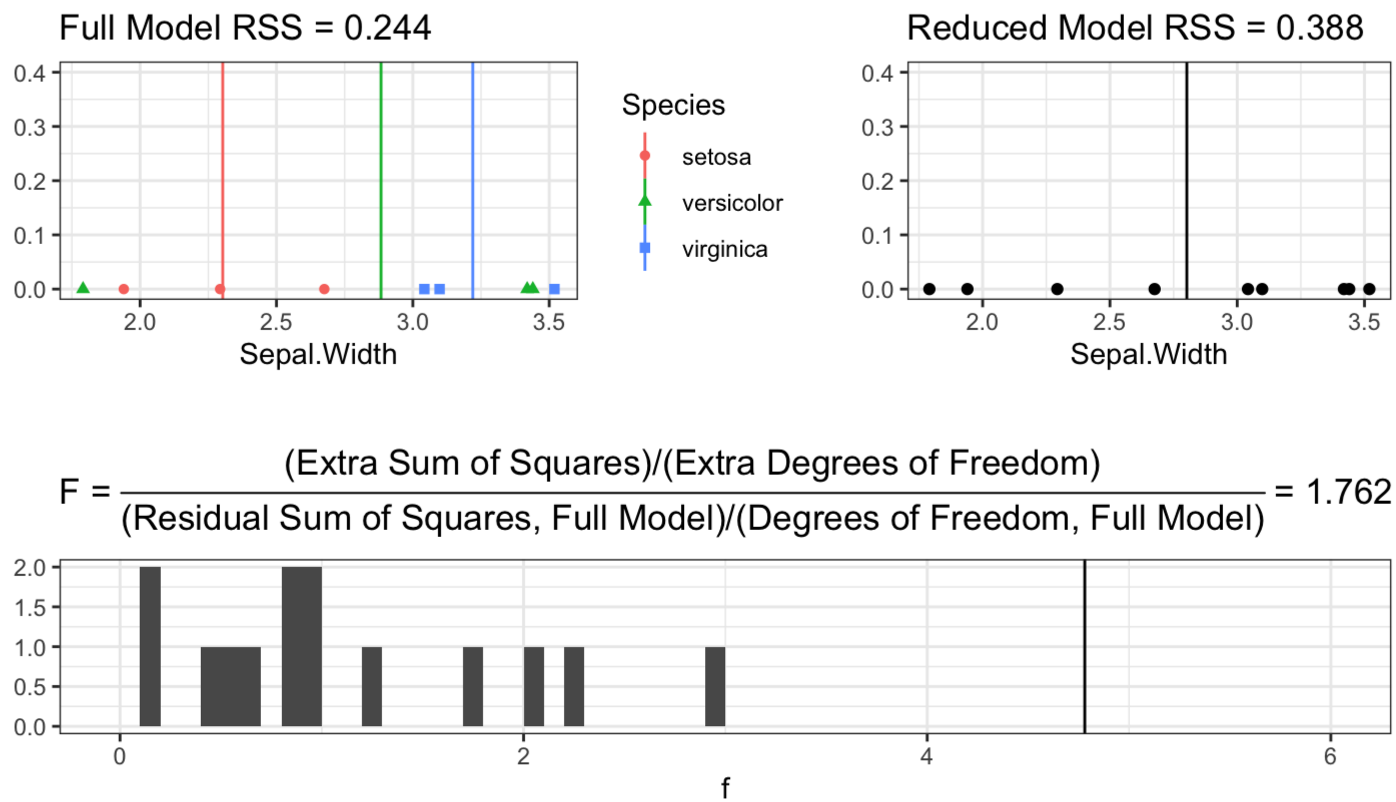
Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



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?

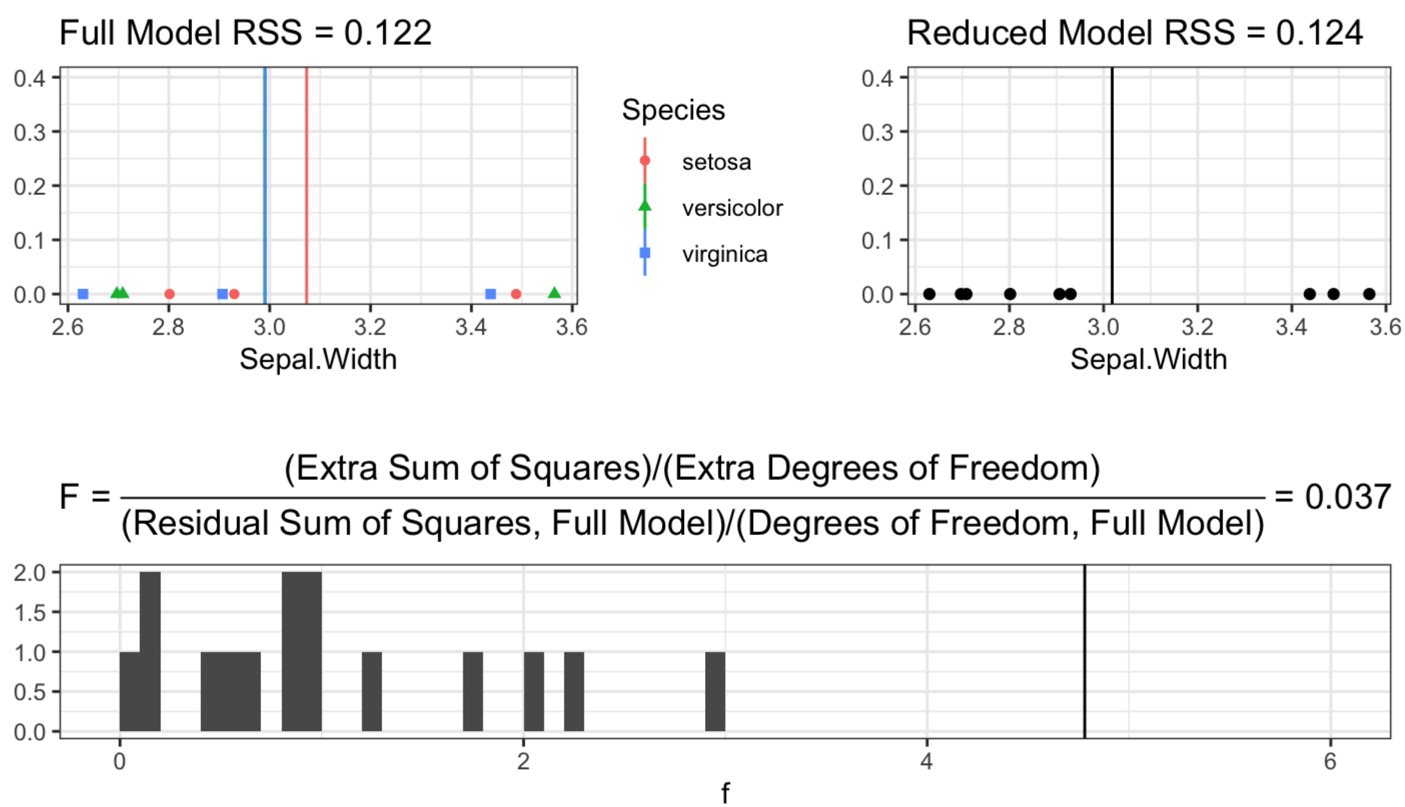
Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



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?

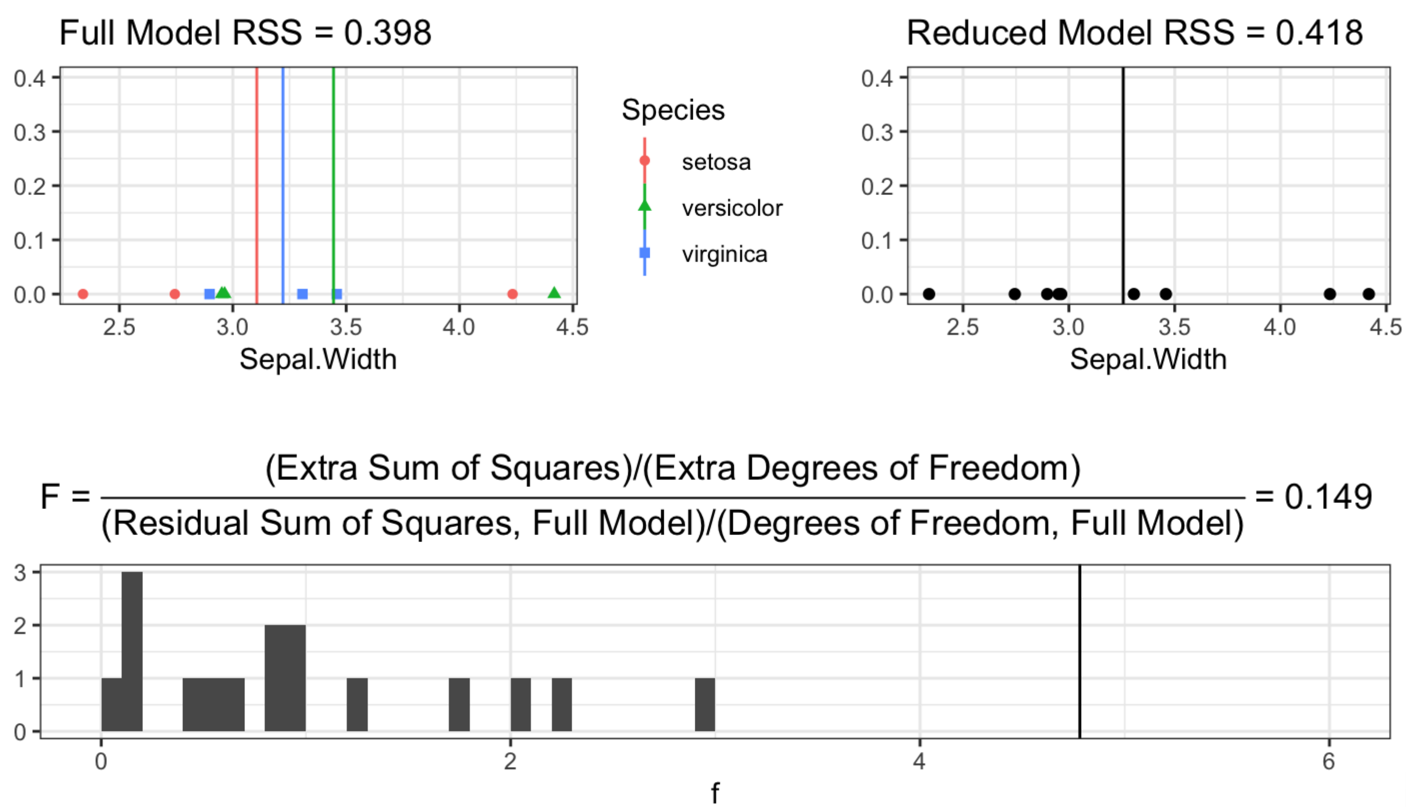
Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



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?

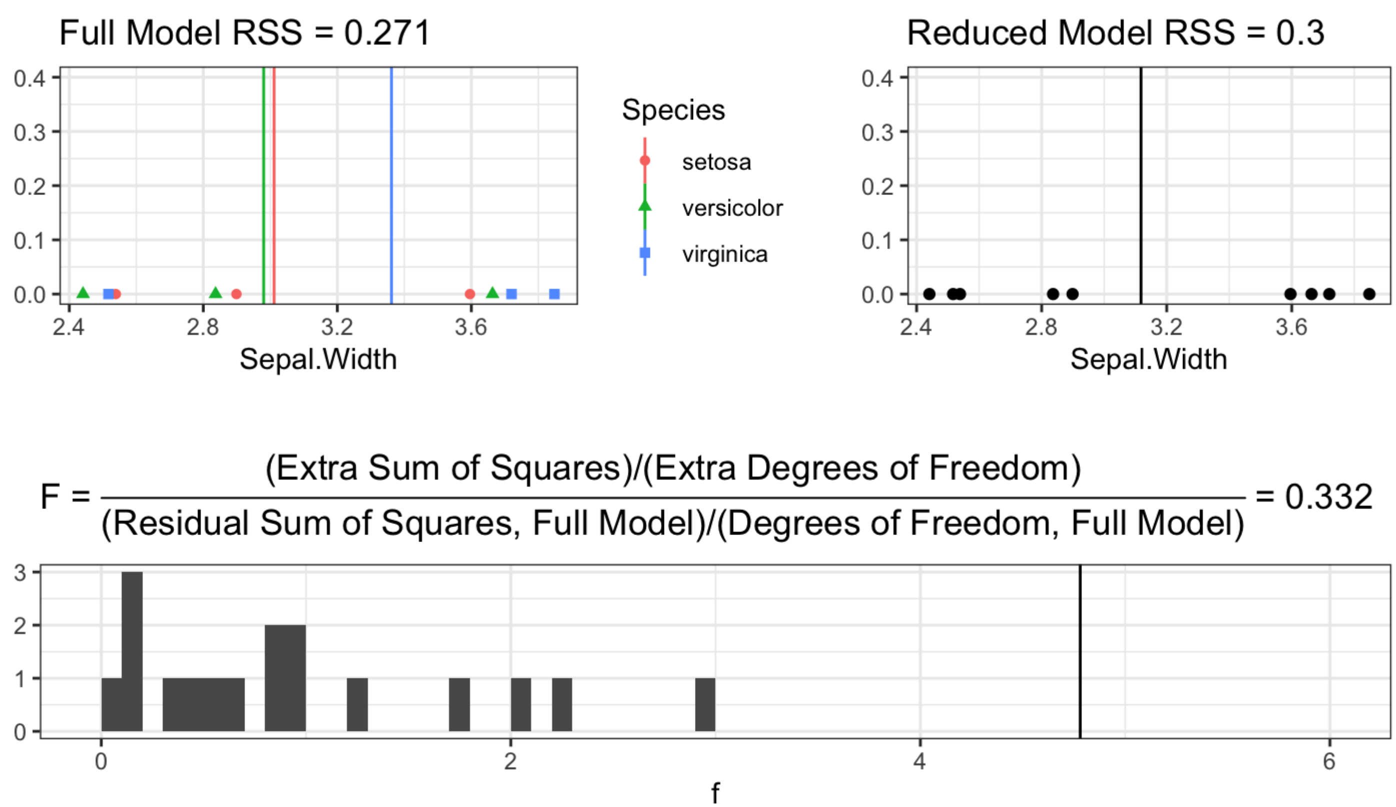
Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



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?

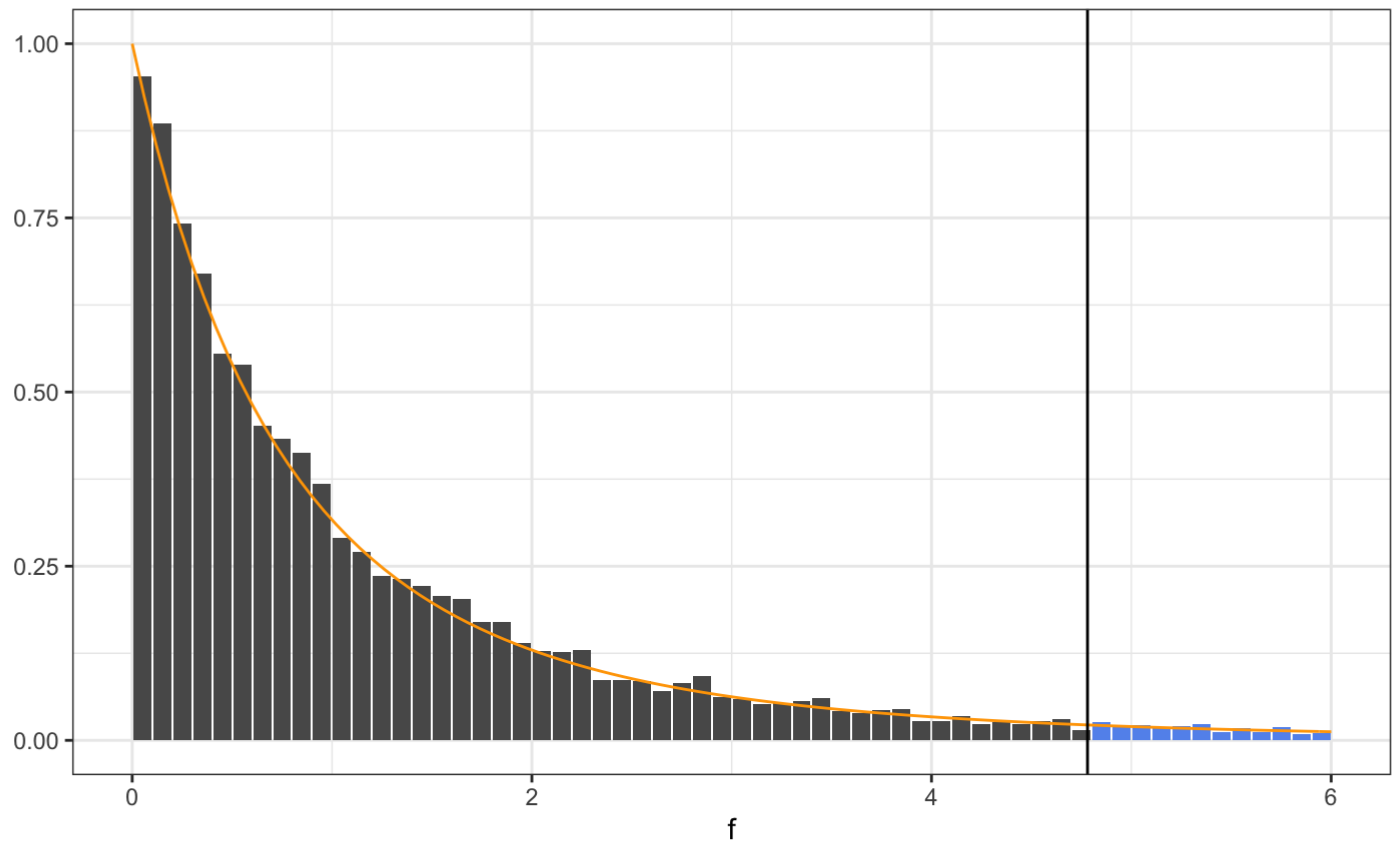
Let's simulate a fake data set, assuming H_0 is true. We have 9 observations drawn at random from a Normal(3.033, 0.579) distribution.



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.