**Stats 201 HW6**

7.1

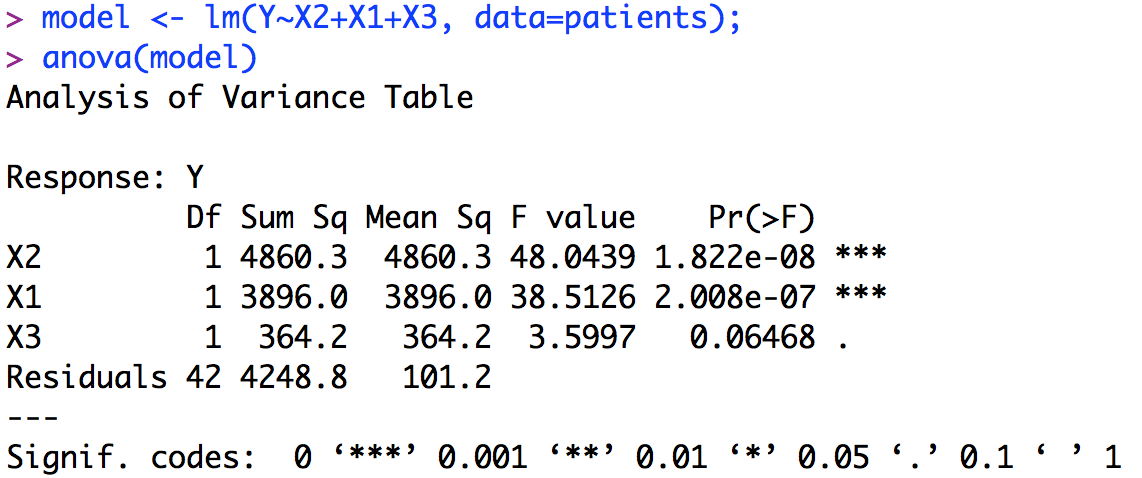
(1) 1 (2) 1 (3) 2 (4) 3

7.2

In the case we don’t use any predictor variables to predict the Y values, we will be simply using as the predicted value (Ŷ), and the SSR will be zero. If we then add X1 in the model to help us predict Y, SSR(X1) will be the extra sum of squares.

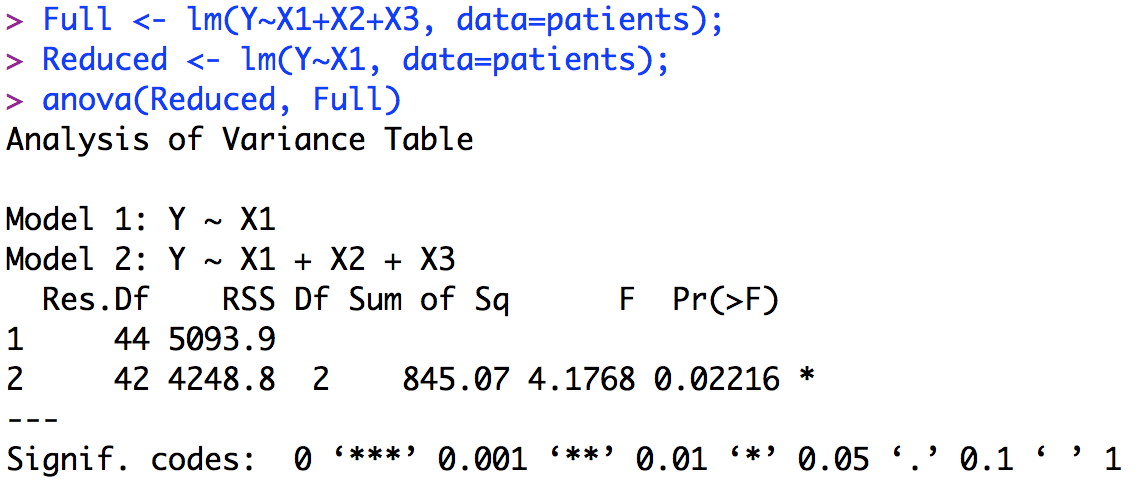
7.5

a. SSR(X2) = 4860.3, SSR(X1|X2) = 3896.0, SSR(X3|X1,X2) = 364.2



b. Null hypothesis: β3 = 0; alternative hypothesis: β3 ≠ 0; decision rule: if p-value < 0.025, reject H0; conclusion: because p-value = 0.0648 > 0.025, accept H0, which means that with X1 and X2 in the model, adding X3 does not help us predict Y.

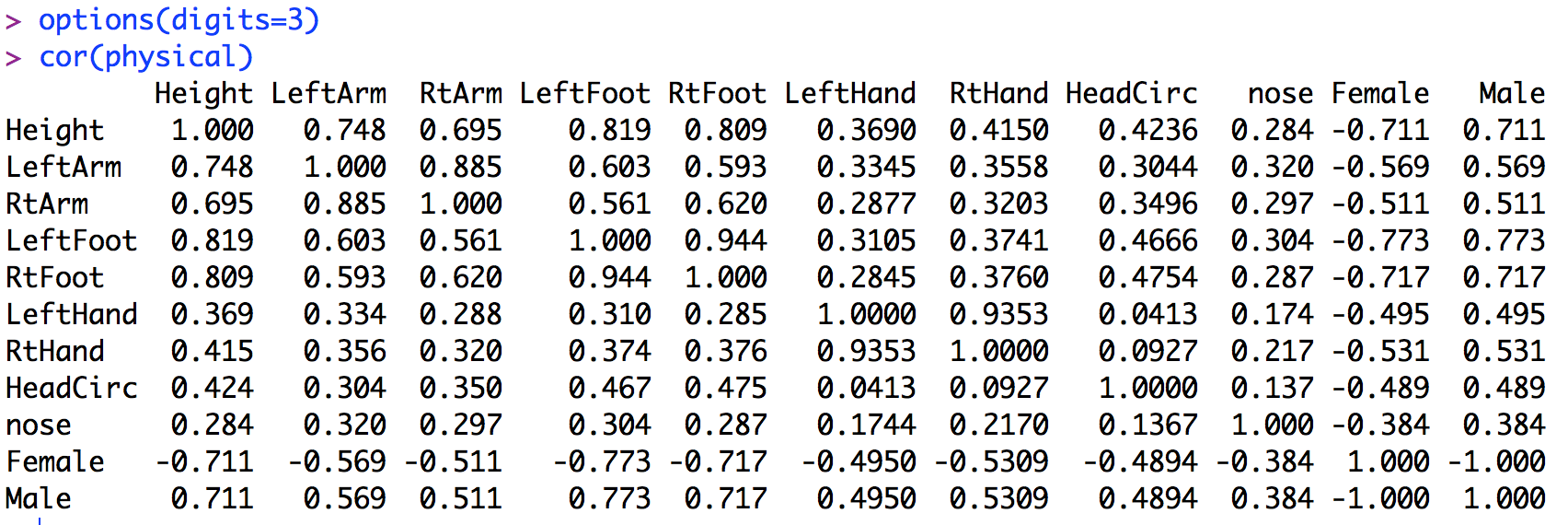
7.6



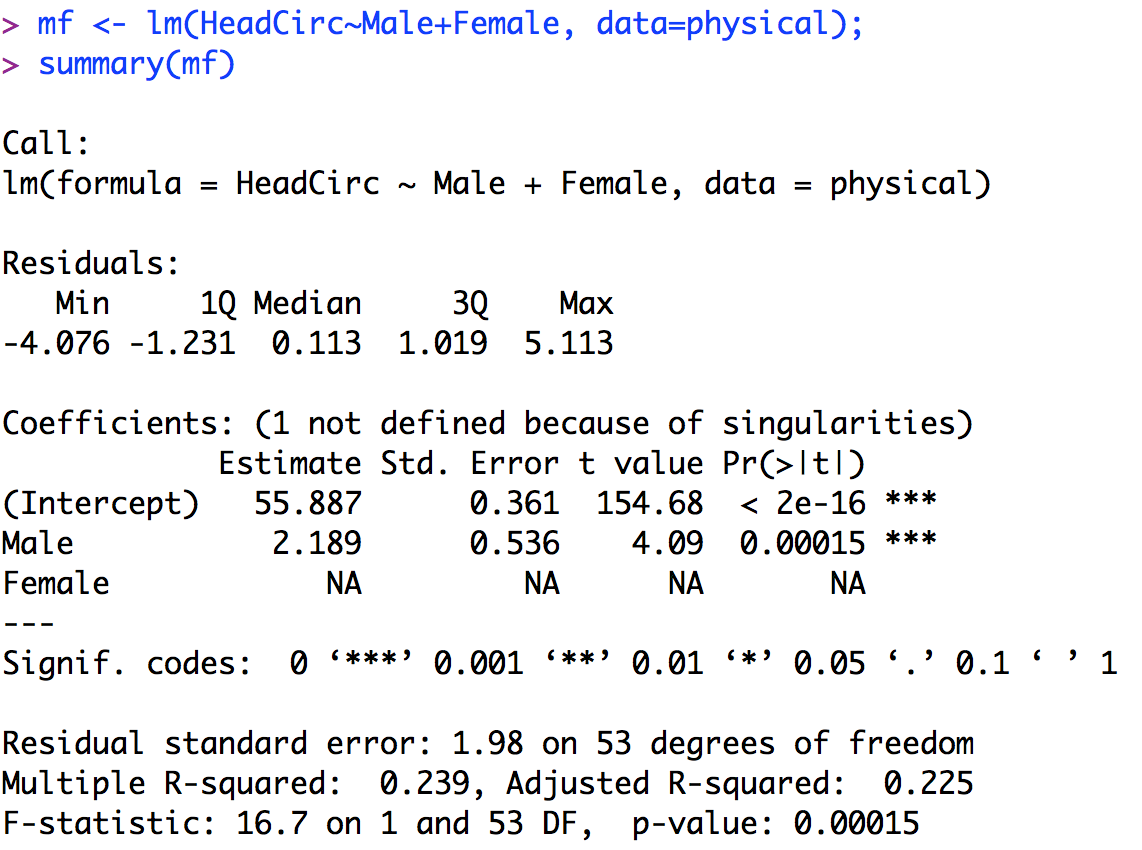
Null hypothesis: β2 = 0, β3 = 0; alternative hypothesis: at least one of β2 and β3 is not zero; decision rule: if p-value < 0.025, reject H0; conclusion: because p-value = 0.02216 < 0.025, reject H0, which means that at least one of X2 and X3 is needed to help us predict Y better.

Answers for the physical data part:

1. cor(Male, Female) = -1, cor(RtFoot, LeftFoot) = 0.944, cor(HeadCirc, RtFoot) = 0.4754, cor(HeadCirc, LeftFoot) = 0.4666



2. R shows that “1 not defined because of singularities”. That is because that Male and Female has perfect correlation (r = -1), so including Female does not help given Male already in the model.



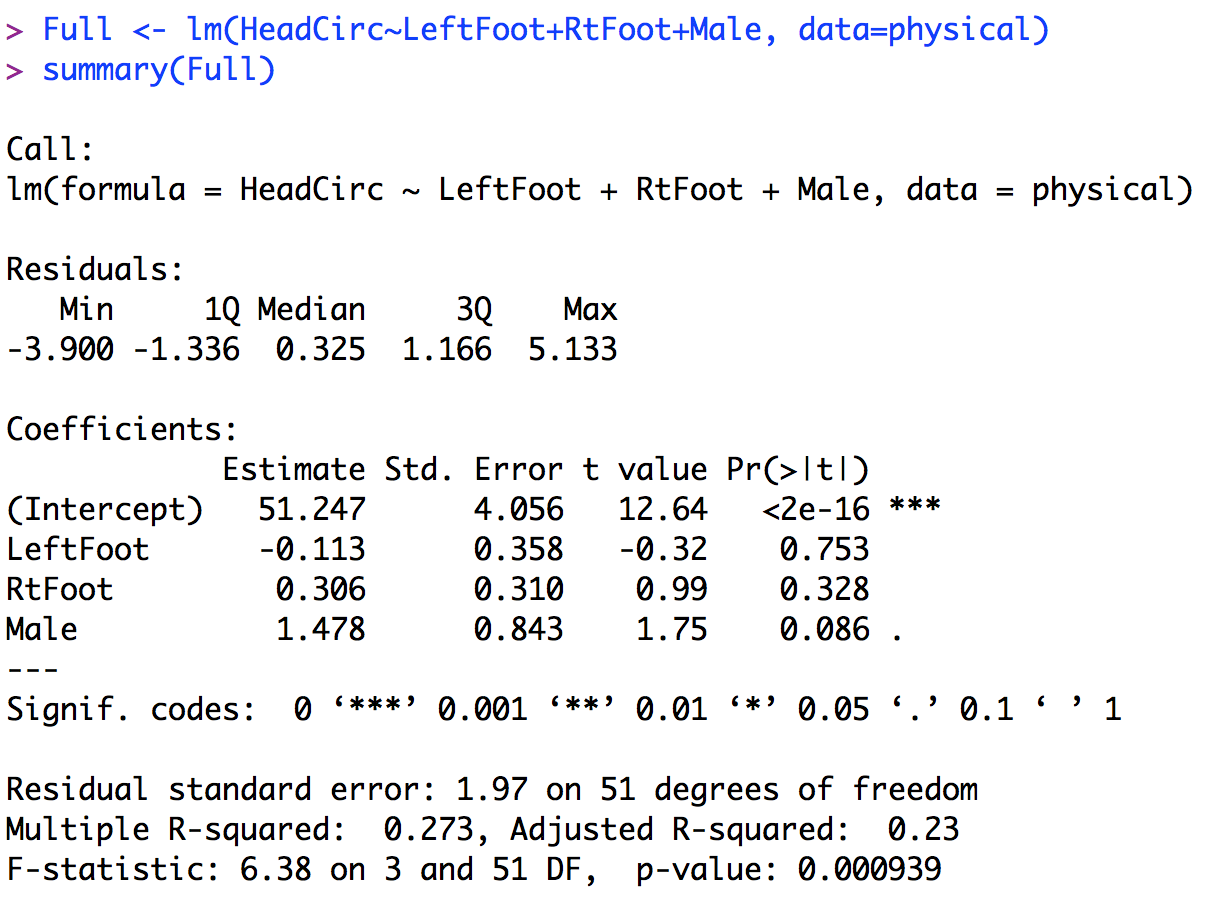
3.

β1: test statistic = -0.32, p-value = 0.753, conclusion: reject Ha, we don’t need LeftFoot given RtFoot and Male in the model.

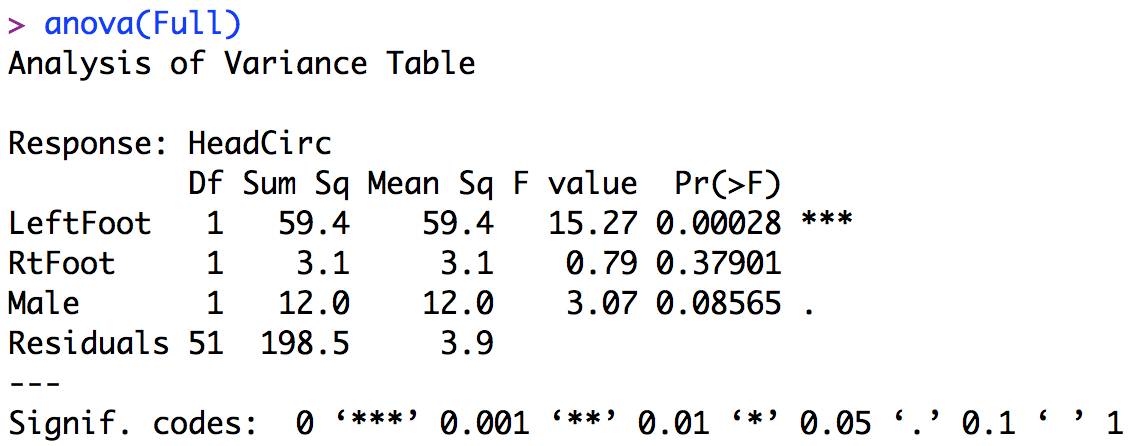
β2: test statistic = 0.99, p-value = 0.328, conclusion: reject Ha, we don’t need RtFoot given LeftFoot and Male in the model.

β3: test statistic = 1.75, p-value = 0.086, conclusion: reject Ha, we don’t need Male given LeftFoot and RtFoot in the model.

None of them are significantly different from zero.



4. p-value = 0.00028 < 0.05, reject H0, conclusion: LeftFoot is needed to predict HeadCirc.

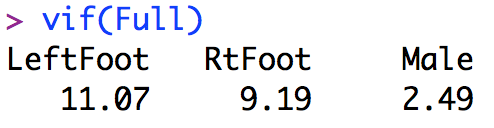


5. In (3), the conclusion is LeftFoot is not needed; in (4), the conclusion is LeftFoot is needed. The difference here is because in (3) we are doing the t-test, so the null hypothesis is LeftFoot is not needed “given RightFoot and Male in the model”. However, in (4), we are doing the F-test and the predictor variables are included in the model one by one. Therefore, the null hypothesis is LeftFoot is not needed “given nothing in the model”. That is why we have two different conclusions.

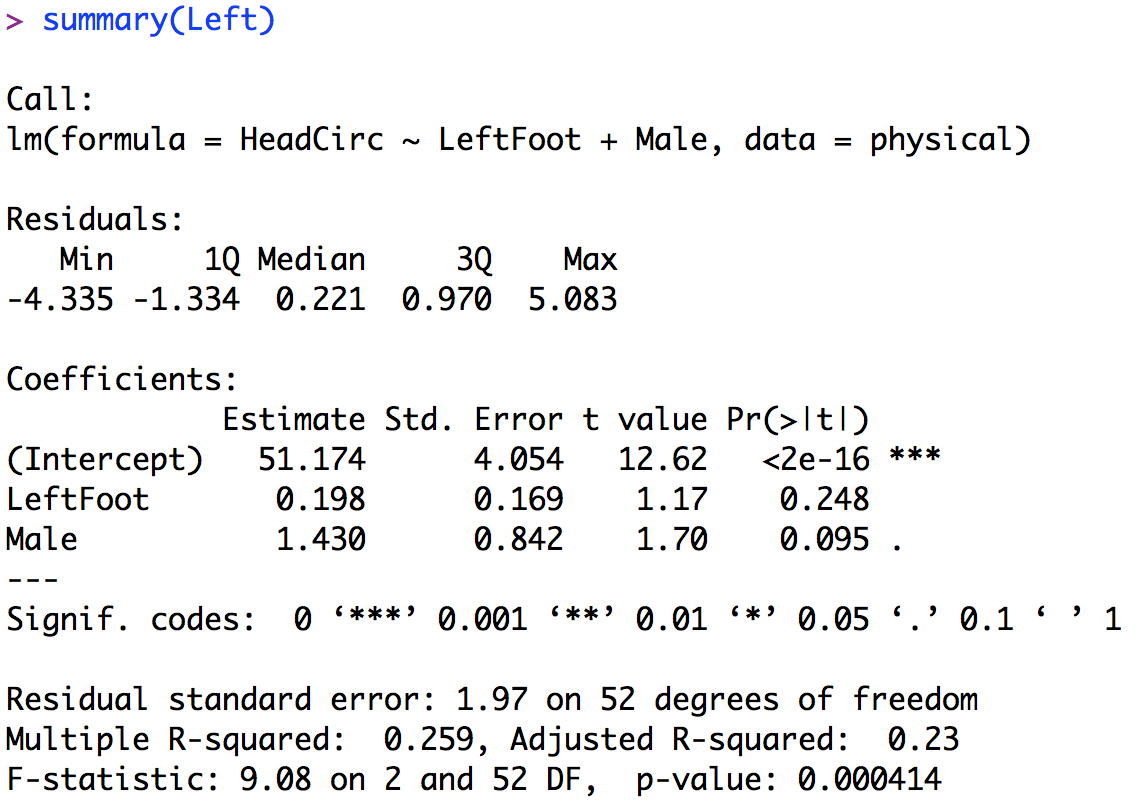
6. From summary(Full), test statistic = 6.38, p-value = 0.000939 < 0.05, so the conclusion is to reject H0, which means at least one of LeftFoot, RightFoot, and Male is needed in the model.

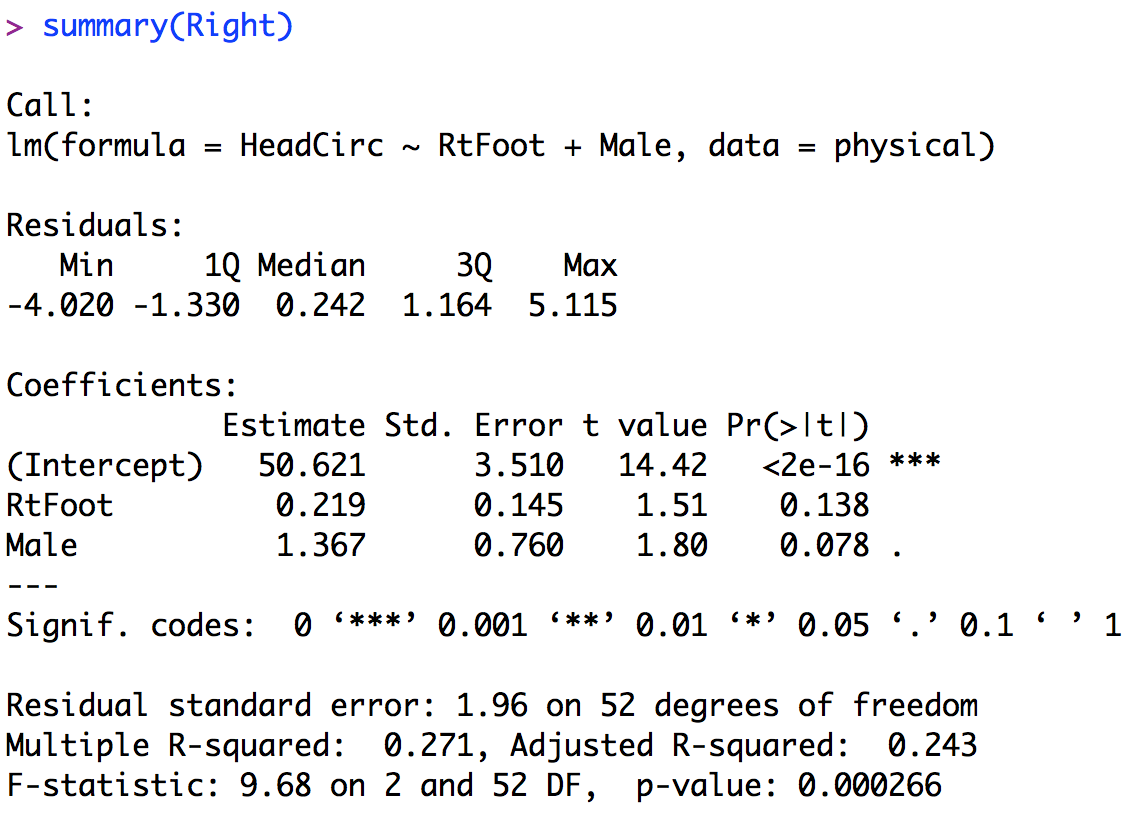
7. The conclusions are slightly different. In (3), the conclusions are that given other two predictor variables in the model, we don’t need an additional variable in the model, so we know we can use any two variable only in the model. However, in (6) the conclusion is that at least one of the variables is needed in the model, but we don’t know which variables are useful.

8. Only include Male in the model, because the VIFs of LeftFoot and RtFoot are greater than 5.



9. In these two models, the p-values of the predictor variables are all less than 0.05. Therefore, we conclude that none of the variables are significant predictors.





10. Right is the best because it has the greatest adjusted R-squared (0.243) and smallest residual standard error (1.96).