Lab 6

Kyle Salitrik
997543474

Part 1

When considering all six predictors, the model 4 predictors that are significant: head length, neck girth, and chest girth. We can tell this from the significantly small p-values calculated from the regression model for these parameters. The head width and body length do not appear to be, when considered individually.

Part 2

If we look at the full model vs the subset of the weight being predicted by only the head width and body length. We obtain the following hypotheses:

$$H_0: \beta_{headwidth} = \beta_{length} = 0$$

$$H_1: \beta_{headwidth} \; ! = 0 \; || \; \beta_{length} \; ! = 0$$

The RSS was calculated from the lm function by obtaining, squaring and summing the residuals to obtain the following values:

$$RSS_{full} = \; 69003.64$$

$$RSS_{reduced} = \; 71386.92$$

As a test statistic, we use the F-value which has the following distribution and p-value:

$$F_{H_0} \sim F_{2,93} = \; 1.606035$$

$$p_{value} = \; 0.2061971$$

Based on the obtained p-value and relative RSS for both models, we can determine that these parameters can potentially be ignored, due to how large the p-value is (p-value >> 0.05).

CODE

```
> bear=read.table('bears.txt',header=TRUE,sep='\t')
> bear=bear[bear$Obs.No==1,]
>
> ## Linear System Solvera
> lm_six = lm(bear$Weight~bear$Head.L+bear$Head.W+bear$Neck.G+bear$Length+bea
r$Chest.G)
> RSS_six = sum(lm_six$residuals^2)
>
>
> lm_reduced = lm(bear$Weight~bear$Head.L+bear$Neck.G+bear$Chest.G)
> RSS_reduced = sum(lm_reduced$residuals^2)
>
>
> ## Set useful variables
> n = dim(bear)[1]
> p=length(lm_six$coefficients)
> q=length(lm_reduced$coefficients)
> f=((RSS_reduced-RSS_six)/(p-q))/(RSS_six/(n-p))
> pvalue=pf(f,p-q,n-p,lower.tail=FALSE)
>
>
> ## Display Results
> summary(lm_six)

Call:
lm(formula = bear$Weight ~ bear$Head.L + bear$Head.W + bear$Neck.G +
    bear$Length + bear$Chest.G)

Residuals:
    Min      1Q  Median      3Q     Max
-59.457 -17.969  -2.059  14.432  99.239

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  -258.3771    20.8837 -12.372  < 2e-16 ***
bear$Head.L    -7.5230     3.3596  -2.239   0.0275 *
bear$Head.W     0.3087     3.3965   0.091   0.9278
bear$Neck.G     8.5812     1.7639   4.865 4.65e-06 ***
bear$Length     1.3305     0.7425   1.792   0.0764 .
bear$Chest.G    7.8844     1.0190   7.738 1.19e-11 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 27.24 on 93 degrees of freedom
Multiple R-squared:  0.9456,   Adjusted R-squared:  0.9427
F-statistic: 323.5 on 5 and 93 DF,  p-value: < 2.2e-16

> summary(lm_reduced)

Call:
lm(formula = bear$Weight ~ bear$Head.L + bear$Neck.G + bear$Chest.G)

Residuals:
    Min      1Q  Median      3Q     Max
```

```
-61.237 -16.783  -2.189  17.310  98.357

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  -247.5019    19.9458 -12.409  < 2e-16 ***
bear$Head.L    -4.1095     2.7264  -1.507    0.135
bear$Neck.G     8.6530     1.6676   5.189 1.20e-06 ***
bear$Chest.G    8.5727     0.9484   9.039 1.87e-14 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 27.41 on 95 degrees of freedom
Multiple R-squared:  0.9438,   Adjusted R-squared:  0.942
F-statistic: 531.3 on 3 and 95 DF,  p-value: < 2.2e-16

> RSS_six
[1] 69003.64
> RSS_reduced
[1] 71386.92
> f
[1] 1.606035
> pvalue
[1] 0.2061971
> anova(lm_reduced,lm_six)
Analysis of Variance Table

Model 1: bear$Weight ~ bear$Head.L + bear$Neck.G + bear$Chest.G
Model 2: bear$Weight ~ bear$Head.L + bear$Head.W + bear$Neck.G + bear$Length
+
    bear$Chest.G
  Res.Df   RSS Df Sum of Sq     F Pr(>F)
1     95 71387
2     93 69004  2    2383.3 1.606 0.2062
```