

Module 5 Lab Submission

Ben Tankus

Again we consider the Brain size data in the data set `case0902` from the `Sleuth3` library. You can read more about this data set by viewing the help file:

```
help(case0902)
```

```
## starting httpd help server ... done
```

```
head(case0902)
```

```
##           Species Brain   Body Gestation Litter
## 1      Aardvark    9.6    2.20         31     5.0
## 2      Acouchis    9.9    0.78         98     1.2
## 3 African elephant 4480.0 2800.00        655     1.0
## 4      Agoutis    20.3    2.80        104     1.3
## 5      Axis deer  219.0   89.00        218     1.0
## 6      Badger    53.0    6.00         60     2.2
```

In Lecture 2 of this Module, we considered modeling all four variables on the log scale, and we compared two models: a full model with log-Body, log-Gestation, and log-Litter as predictors, and a reduced model with only log-Body as a predictor.

1. Fit the two models described above (and as shown on slide 15 of Lecture 2).

```
fitFull <- lm(log(Brain) ~ log(Body) + log(Gestation) + log(Litter), data = case0902)
fitRed <- lm(log(Brain) ~ log(Body), data = case0902)
```

```
#summary(fitFull)
#summary(fitRed)
```

Now suppose we have four new species, with the following values of Body, Gestation, and Litter:

Species	Body	Gestation	Litter
A	1	50	1
B	1	50	2
C	1	150	1
D	50	50	1

2. Create a new data frame with these values, and use the `predict()` function to predict the log-brain size for these new species.

```
newData = data.frame(
  Species = c('A', 'B', 'C', 'D')
, Body = c(1,1,1,50)
, Gestation = c(50,50,150,50)
, Litter = c(1,2,1,1)
)
print('FULL MODEL')
```

```
## [1] "FULL MODEL"
```

```
predict(fitFull, newdata = newData)
```

```
##          1          2          3          4
## 2.489821 2.274896 2.948977 4.739513
```

```
print('REDUCED MODEL')
```

```
## [1] "REDUCED MODEL"
```

```
predict(fitRed, newdata = newData)
```

```
##          1          2          3          4
## 2.332351 2.332351 2.332351 5.145825
```

3. Obtain confidence intervals for the mean log-brain size of the new species above, using the `interval = "confidence"` argument in the `predict()` function.

```
print('FULL MODEL')
```

```
## [1] "FULL MODEL"
```

```
predict(fitFull, newdata = newData, interval = 'confidence')
```

```
##          fit          lwr          upr
## 1 2.489821 2.208628 2.771014
## 2 2.274896 2.116014 2.433778
## 3 2.948977 2.732987 3.164968
## 4 4.739513 4.312767 5.166260
```

```
print('REDUCED MODEL')
```

```
## [1] "REDUCED MODEL"
```

```
predict(fitRed, newdata = newData, interval = 'confidence')
```

```
##          fit          lwr          upr
## 1 2.332351 2.186921 2.477782
## 2 2.332351 2.186921 2.477782
## 3 2.332351 2.186921 2.477782
## 4 5.145825 5.008289 5.283360
```

4. Obtain prediction intervals for the log-brain size of the new species above, using the `interval = "prediction"` argument in the `predict()` function.

```
print('FULL MODEL')
```

```
## [1] "FULL MODEL"
```

```
predict(fitFull, newdata = newData, interval = 'prediction')
```

```
##          fit          lwr          upr
## 1 2.489821 1.505881 3.473761
## 2 2.274896 1.318700 3.231092
## 3 2.948977 1.981651 3.916303
## 4 4.739513 3.704535 5.774492
```

```
print('REDUCED MODEL')
```

```
## [1] "REDUCED MODEL"
```

```
predict(fitRed, newdata = newData, interval = 'prediction')
```

```
##          fit          lwr          upr
## 1 2.332351 1.175408 3.489294
## 2 2.332351 1.175408 3.489294
## 3 2.332351 1.175408 3.489294
## 4 5.145825 3.989847 6.301802
```

5. Compare the reduced model to the full model using the AIC and BIC measures.

```
anova(fitFull, fitRed) #ANOVA SHOWS SIGNIFICANT DIFFERENCE
```

```
## Analysis of Variance Table
##
## Model 1: log(Brain) ~ log(Body) + log(Gestation) + log(Litter)
## Model 2: log(Brain) ~ log(Body)
##   Res.Df    RSS Df Sum of Sq    F    Pr(>F)
## 1      92 20.736
## 2      94 31.411 -2   -10.675 23.681 5.053e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
AIC(fitFull, fitRed) #AIC
```

```
##          df          AIC
## fitFull   5 135.3188
## fitRed    3 171.1866
```

```
BIC(fitFull,fitRed)
```

```
##           df      BIC
## fitFull  5 148.1405
## fitRed   3 178.8796
```

In both AIC and BIC the full model is superior and has a lower value than the reduced model.