

DVR Weight-Length Relation

Preliminaries

Load Necessary Packages

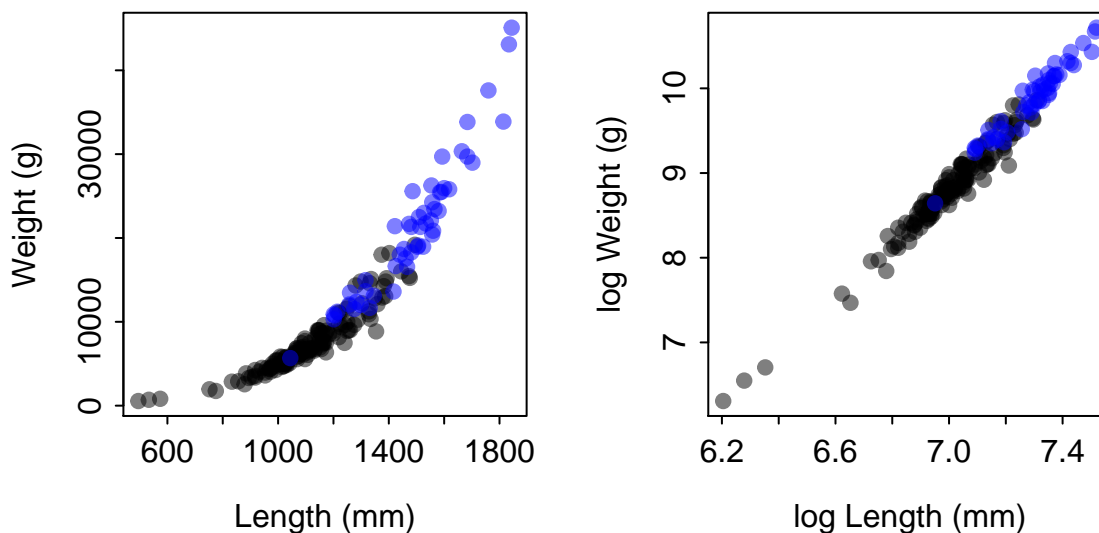
```
> library(FSA)      # for filterD(), hist(), Summarize()
> library(dplyr)     # for mutate(), select()
```

Load Data

```
> # Set your working directory to where your external data files (and scripts) are located.
> setwd("C:/aaaWork/Web/GitHub/RcourseNunavut2016/Handouts")
> dSC <- read.csv("SawyerCo_reduced.csv")
> dSC <- mutate(dSC, loglen=log(len), logwt=log(weight))
> Sturg <- filterD(dSC, waterbody %in% c("CHIPPEWA RIVER", "HUNTER LAKE"),
  species=="Lake Sturgeon", !is.na(len), !is.na(weight))
```

Quick Summaries

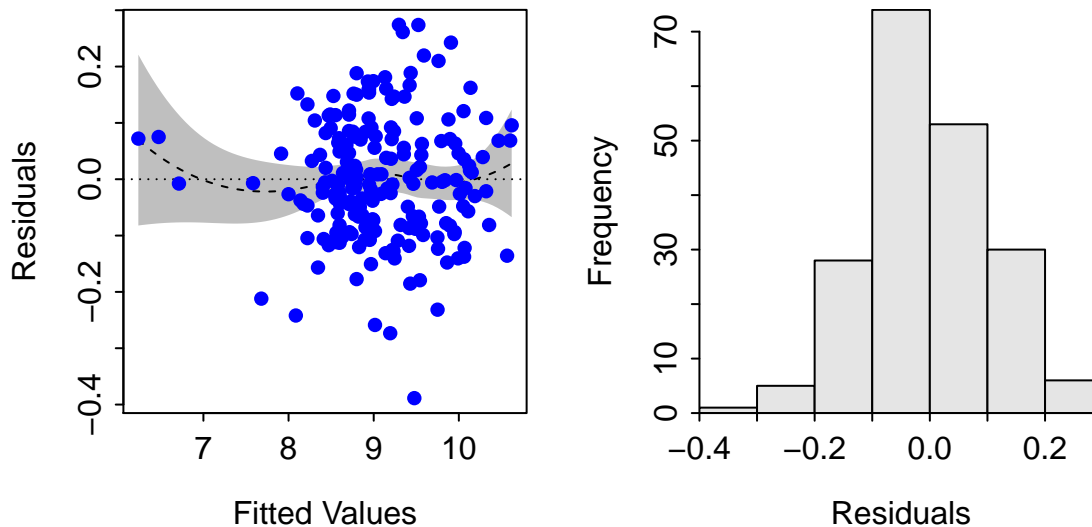
```
> clr1 <- c("black", "blue")
> clr2 <- col2rgbt(clr1, 1/2)
> plot(weight~len, data=Sturg, pch=19, col=clr2[waterbody],
  xlab="Length (mm)", ylab="Weight (g)" # Left
> plot(logwt~loglen, data=Sturg, pch=19, col=clr2[waterbody],
  xlab="log Length (mm)", ylab="log Weight (g)" # Right
```



Dummy Variable Regression

Fitting the Model

```
> dvr1 <- lm(logwt~loglen*waterbody,data=Sturg)  
> residPlot(dvr1,legend=FALSE)
```



```
> anova(dvr1)  
Analysis of Variance Table  
  
Response: logwt  
          Df Sum Sq Mean Sq  F value    Pr(>F)  
loglen      1  94.398   94.398 7879.7860 < 2.2e-16  
waterbody   1   0.356    0.356  29.7194 1.514e-07  
loglen:waterbody 1   0.004    0.004   0.3712  0.5431  
Residuals 193   2.312    0.012
```

```
> dvr2 <- lm(logwt~loglen+waterbody,data=Sturg)  
> anova(dvr2)  
Analysis of Variance Table  
  
Response: logwt  
          Df Sum Sq Mean Sq  F value    Pr(>F)  
loglen      1  94.398   94.398 7905.408 < 2.2e-16  
waterbody   1   0.356    0.356  29.816 1.443e-07  
Residuals 194   2.317    0.012
```

```
> summary(dvr2)  
  
Call:  
lm(formula = logwt ~ loglen + waterbody, data = Sturg)  
  
Residuals:  
      Min       1Q   Median       3Q      Max   
-0.39137 -0.07853 -0.00828  0.07521  0.27213  
  
Coefficients:
```

```

              Estimate Std. Error t value Pr(>|t|)
(Intercept)    -13.84283    0.34946  -39.61  < 2e-16
loglen          3.23442    0.04991   64.80  < 2e-16
waterbodyHUNTER LAKE  0.12556    0.02299    5.46 1.44e-07

Residual standard error: 0.1093 on 194 degrees of freedom
Multiple R-squared:  0.9761,    Adjusted R-squared:  0.9759
F-statistic: 3968 on 2 and 194 DF,  p-value: < 2.2e-16

```

```

> round(cbind(ests=coef(dvr2),confint(dvr2)),3)
              ests      2.5 %   97.5 %
(Intercept)    -13.843  -14.532  -13.154
loglen          3.234   3.136   3.333
waterbodyHUNTER LAKE  0.126   0.080   0.171

```

Making Predictions

```

> tmp <- data.frame(loglen=log(c(1000,1000)),waterbody=c("CHIPPEWA RIVER","HUNTER LAKE"))
> predict(dvr2,tmp,interval="confidence")
      fit      lwr      upr
1 8.499773 8.479738 8.519809
2 8.625331 8.576953 8.673708

```

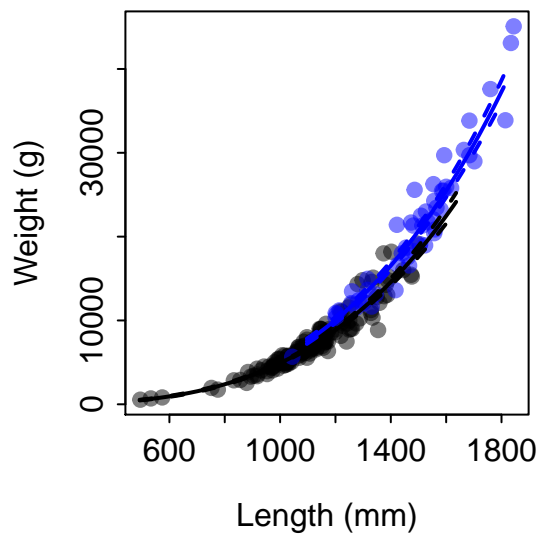
Summary Plot

```

> Summarize(loglen~waterbody,data=Sturg,digits=1)
  waterbody  n nvalid mean  sd min  Q1 median  Q3 max percZero
1 CHIPPEWA RIVER 145    145  7.0 0.2 6.2 6.9   7.0 7.1 7.3      0
2  HUNTER LAKE  52     52  7.3 0.1 7.0 7.2   7.3 7.4 7.5      0

> logL <- seq(6.2,7.4,length.out=199)
> logW <- predict(dvr2,data.frame(loglen=logL,waterbody="CHIPPEWA RIVER"),interval="confidence")
> cL <- exp(logL)
> cW <- exp(logW)
> logL <- seq(7.0,7.5,length.out=199)
> logW <- predict(dvr2,data.frame(loglen=logL,waterbody="HUNTER LAKE"),interval="confidence")
> hL <- exp(logL)
> hW <- exp(logW)
> plot(weight~len,data=Sturg,pch=19,col=clr2[waterbody],
      xlab="Length (mm)",ylab="Weight (g)")
> lines(cL,cW[, "fit"],lwd=2,col=clr1[1])
> lines(cL,cW[, "lwr"],lwd=2,lty=2,col=clr1[1])
> lines(cL,cW[, "upr"],lwd=2,lty=2,col=clr1[1])
> lines(hL,hW[, "fit"],lwd=2,col=clr1[2])
> lines(hL,hW[, "lwr"],lwd=2,lty=2,col=clr1[2])
> lines(hL,hW[, "upr"],lwd=2,lty=2,col=clr1[2])

```



Different Summary

```
> lwCompPreds(dvr2)
```

