1. True

2. True

3. False

4. False

---

Model Building Qn 1:

1. choose one w/ smallest AIC

   Weight ~ Length3

2. Weight ~ Length3 + Width

3. $R^2_{adj} = 1 - \frac{(n-1)}{(n-3)} \times \frac{SSE}{SST} = 1 - \frac{158}{156} \times \frac{2489410}{20247459}$

   $= 0.8755$

4. $C_p = \frac{SSE_p}{MSE_F} + 2p' - n$

   $= \frac{2489410}{12.66^2} + 6 - 159$

   $= 15379.058$

   $AIC = n\log(SSE) - n\log(n) + 2p'$

   $1541.7 = 159\log(SSE) - 159\log(15) + 6$

   $SSE = 2489007.937$

   $C_p$ is much less than $p'$, model is too simple a full set of predictors is needed

5. qt(0.975, df = 156)

   Std resid ~ t(df = n-p-1)

Model Q2

1.

$n\log(SSE) - n\log(n) + 2p'$

$= 18\log(0.11517) - 18\log(18) + 2(5)$

$= -80.93042$

2.

$n\log(SSE) - n\log(n) + 2p'$

$n = 18$

$2p' = 8$

$C_p = 3.086 = \frac{SSE_p}{MSE_{Full}} + 2p' - n$

$\Rightarrow SSE_p = (10 + 3.086) \times 0.00886$

$= 0.115942$

$18\log(0.115942) - 18\log(18) + 8 = -82.811$

3. Model 2 is better since AIC is lower

4. $R^2_{adj} = 1 - \frac{(n-1)}{(n-2)} \times \frac{SSE_1}{SST}$

$C_p = 19.176 = \frac{SSE_1}{MSE_F} + 2p' - n$

$\Rightarrow SSE_1 = 0.2439394$

$SST = 0.55467 + 0.11577 + 0.02761 + 0.03336 + 0.11517$

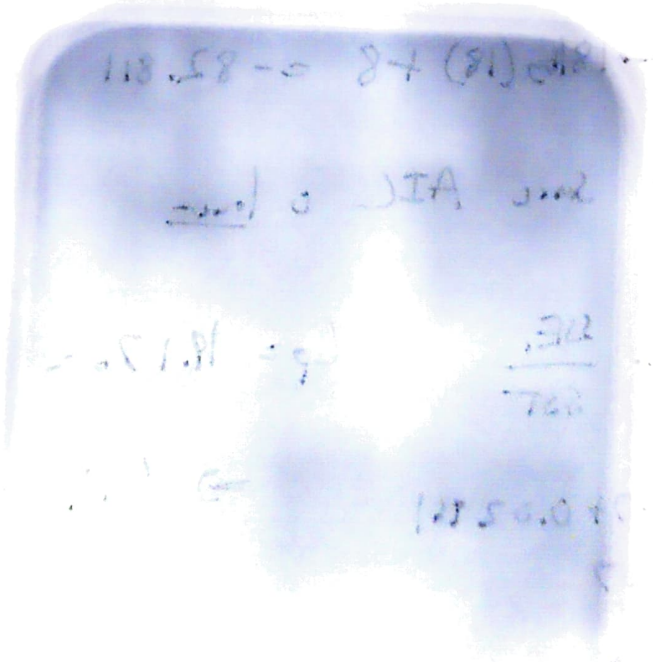
$= 0.84858$

$R^2_{adj} = 1 - \frac{17}{16} \times \frac{0.293\ldots}{0.848\ldots} = 0.632$

5. $VIF = \frac{1}{1-r^2}$

$= \frac{1}{1-(0.84)^2} = 3.3467$

not too worried



# Forward Selection

$M_0: y = \beta_0 + \varepsilon$

$y = \beta_0 + \beta_j x_j + \varepsilon \qquad j = 1, \ldots, p$

- choose the predictor that gives largest dip in SSE

$$SS_{YY} = \sum_i (y_i - \bar{y})^2 \qquad SS_{xx}^{(j)} = \sum_i (x_{ij} - \bar{x}_j)^2 \qquad SS_{xy}^{(j)} = \sum_i (x_{ij} - \bar{x}_j)(y_i - \bar{y})$$

$$\hat{\beta}_j = \frac{SS_{xy}^{(j)}}{SS_{xx}^{(j)}} \qquad b = \bar{y} - \hat{B}\bar{x}$$

$$SSE = SS_{YY} - \frac{(SS_{xy}^{(j)})^2}{SS_{xx}^{(j)}}$$

$$SSE = \sum (y_i - (\hat{B}x_i + b))$$
$$= \sum (y_i - \hat{B}x_i - \bar{y} + \hat{B}\bar{x})^2$$
$$= \sum [\hat{B}(\bar{x} - x_i) - (\bar{y} - y_i)]^2$$
$$= \hat{B}^2 SS_{xx} - 2\hat{B} SS_{xy} + SS_{YY}$$
$$= SS_{YY} - \hat{B} SS_{xy}$$
$$\hat{B} = \frac{SS_{xy}}{SS_{xx}}$$

$$r_j = \frac{SS_{xy}^{(j)}}{\sqrt{SS_{xx}^{(j)} SS_{YY}}} \qquad r = \frac{\mathrm{cov}(x,y)}{\sigma_x \sigma_y}$$

$$\frac{(SS_{xy}^{(j)})^2}{SS_{xx}^{(j)}} = r_j^2 SS_{YY}$$

$$SSE = SS_{YY} - r_j^2 SS_{YY}$$
$$= SS_{YY}(1 - r_j^2)$$

$SS_{YY}$ does not depend on $j$, minimizing SSE is the same as maximizing $r_j^2$ or rather maximizing corr.

$$AIC = n \log(SSE) - n \log(n) + 2p'$$

~~$AIC = n \log(SSE) - n \log(n) + 2p'$~~

min AIC is equiv. to minimizing RSS

# LS Q 1

$$Var(\varepsilon_i) = \sigma^2 X_i$$

$$Y_i = B_1 X_i + \varepsilon_i$$

$$\frac{Y_i}{\sqrt{X_i}} = B_1 \frac{X_i}{\sqrt{X_i}} + \frac{\varepsilon_i}{\sqrt{X_i}} = B_1 \sqrt{X_i} + \varepsilon_i^*$$

$$Y_i^* = \frac{Y_i}{\sqrt{X_i}} \qquad X_i^* = \sqrt{X_i} \qquad \varepsilon_i^* = \frac{\varepsilon_i}{\sqrt{X_i}}$$

$$Y_i^* = B_1 X_i^* + \varepsilon_i^*$$

$$Var(\varepsilon_i) = \frac{Var(\varepsilon_i)}{\sqrt{X_i}} = \frac{\sigma^2 X_i}{X_i} = \sigma^2 \quad \text{not constant}$$

$$\hat{\beta}(\lambda) = (X^T X + \lambda I)^{-1} X^T y$$

- since we only have one regressor, $X^T X$, $X^T y$ are scalars

$$X^T X = \sum x_i^2 = (-1)^2 + 1^2 + 1^2 + (-1)^2 = 4$$

$$X^T y = \sum x_i y_i = 1.5 + 2.9 - 3.5 - 0.7 \approx 0.2$$

$$\hat{\beta}(\lambda) = \frac{X^T y}{X^T X + \lambda} = \frac{0.2}{4+\lambda} = \frac{1}{5(4+\lambda)}$$

a)

$$\hat{\beta}(0) = \frac{0.2}{4+0} = \frac{1}{20} = 0.05$$

b)

$$\hat{\beta}(1) = \frac{0.2}{4+1} = \frac{0.2}{5} = \frac{1}{25} = 0.04$$

c).

$$\hat{\beta}(\lambda) = \frac{0.2}{4+\lambda}$$

$|\hat{\beta}(\lambda)|$ decreases monotonically as $\lambda$ increases

$\hat{\beta}(\lambda) \to 0$ as $\lambda \to \infty$