In these notes we are going to compare three models that we have studied this semester:

- Equal mean model (i.e., a 1-sample approach)
- A SLR
- ANOVA (i.e., separate means model)

We are going to go over the mechanics for performing an extra sum of squares lack of fit test to assist us in determining which model is “best” for describing the data.

Some housekeeping:

```r
library(Sleuth3)
```

```r
## Warning: package 'Sleuth3' was built under R version 3.3.3

source("http://www.math.montana.edu/parker/courses/STAT411/diagANOVA.r")
```

### 8.1 Case study of breakdown times under different voltages

Here’s the data for the case study in section 8.1.2 where ‘breakdown time’ of an insulating fluid was studied in an experiment under uniform conditions. Let’s fit the SLR

\[
\mu\{\text{Time}|\text{Voltage}\} = \beta_0 + \beta_1 \text{Voltage}
\]

```r
# Get data and plot it
summary(case0802)
```

```r
## Time Voltage Group
## Min. : 0.090 Min. :26.00 Group1: 3
## 1st Qu.: 1.617 1st Qu.:31.50 Group2: 5
## Median : 6.925 Median :34.00 Group3:11
## Mean : 98.558 Mean :33.13 Group4:15
## 3rd Qu.: 38.383 3rd Qu.:36.00 Group5:19
## Max. :2323.700 Max. :38.00 Group6:15
## Group7: 8
```  

```r
dim(case0802)    # This tells us that n=76
```

```r
## [1] 76 3
```

```r
m.BAD=lm(Time ~ Voltage,data=case0802)
```

```r
plot(Time ~ Voltage,data=case0802)
abline(coef(m.BAD))    # Show the line
```

```r
volt=seq(26,38,length=100)
lines(volt,exp(18.96 -0.507*volt),col="red",lty=2)    # We'll see below where this comes from
legend("topright",legend = c("linear","exponential"),bty = "n",col = c("black","red"), lty =c(1,2))
```
diagANOVA(m.BAD)
The scatterplot and residual vs. fits plot show that Time and Voltage do not have a linear relationship. Your book suggests what to do: log-transform the response! The new SLR model is the negative exponential

\[
\text{Median}\{\text{Time}|\text{Voltage}\} = e^{\beta_0 + \beta_1 \text{Voltage}}
\]

```
m.SLR = lm(log(Time) ~ Voltage, data = case0802)
summary(m.SLR)
```

```R
## Call:
## lm(formula = log(Time) ~ Voltage, data = case0802)
##
## Residuals:
```
## Min 1Q Median 3Q Max
## -4.0291 -0.6919 0.0366 1.2094 2.6513

## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 18.9555 1.9100 9.924 3.05e-15 ***
## Voltage -0.5074 0.0574 -8.840 3.34e-13 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Residual standard error: 1.56 on 74 degrees of freedom
## Multiple R-squared: 0.5136, Adjusted R-squared: 0.507
## F-statistic: 78.14 on 1 and 74 DF, p-value: 3.34e-13

```r
plot(log(Time) ~ Voltage, data=case0802)
abline(coef(m.SLR))
```

![Graph showing log(Time) vs Voltage]

diagANOVA(m.SLR)
The scatterplot and residual vs. fits plot for an SLR of log(Time) vs Voltage indicate much better fit to the data.

We have not looked at an ANOVA table for SLR before. Compare with Display 8.8.

```r
anova(m.SLR)
```

```
## Analysis of Variance Table
##
## Response: log(Time)
## Df Sum Sq Mean Sq F value Pr(>F)
## Voltage 1 190.15 190.151 78.141 3.34e-13 ***
## Residuals 74 180.07 2.433
## ---
```
The ANOVA table allows us to use sums of squares to confirm that $R^2 = 51\%$ (cf. `summary(m.SLR)` output above). The formula $R^2 = 1 - \frac{\text{ResidualSS}}{\text{TotalSS}}$ was provided in the Chapter 5 notes:

\[
1 - \frac{180.07}{(180.07 + 190.15)}
\]

Because only 7 voltages were tested in this experiment, we can also consider a ANOVA fit to these data.

```r
m.ANOVA = lm(log(Time) ~ as.factor(Voltage), data=case0802)
anova(m.ANOVA)  # Compare with Display 8.8.
```

The last model we will need is the “equal means” model. We can fit it like this:

```r
m.null = lm(log(Time) ~ 1, data=case0802)
anova(m.null)
```

## Analysis of Variance Table
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

1 - \frac{173.75}{(173.75 + 196.48)}  # can also check via `summary(m.ANOVA)`

But, remember, we discussed that $R^2$ always increases as the number of parameters increases. How many parameters are in each of the models above?

### 8.5 Extra sum of squares lack of fit test

The extra sum of squares lack of fit test can help us to decide whether the equal means model sufficiently describes the data; or whether the SLR “better” describes it; or whether the ANOVA is better.

The 6 steps when comparing SLR to ANOVA test:

1. $H_0$: SLR is the true model vs. $H_a$: ANOVA is the true model
2. Check the assumptions for SLR and ANOVA

3. Test statistic is \( F_{stat} = \frac{SS_{SLR} - SS_{ANOVA}}{SS_{ANOVA}/DF_{ANOVA}} \) where \( SS_{SLR} \) and \( SS_{ANOVA} \) are the Residual sums of squares from the ANOVA tables above for the SLR and ANOVA models respectively. So \( F_{stat} \) is a measure of the drop in the Residual sums of squares between the two models. This drop must be large to conclude that the additional parameters in the ANOVA model substantially improve model fit.

4. \( p \)-value is \( P(F > F_{stat}) \) where \( F \sim F(I, n - I) \)

5. Make a decision

6. State a conclusion - a small \( p \)-value leads you to conclude that the model under \( H_a \) is "better"

It is easy to implement this test in R

```r
anova(m.null,m.SLR,m.ANOVA)
```

## Analysis of Variance Table
##
## Model 1: log(Time) ~ 1
## Model 2: log(Time) ~ Voltage
## Model 3: log(Time) ~ as.factor(Voltage)
##
## Res.Df RSS Df Sum of Sq F Pr(>F)
## 1 75 370.23
## 2 74 180.07 1 190.151 75.5139 1.096e-12 ***
## 3 69 173.75 5 6.326 0.5024 0.7734
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

\[
\frac{6.326/5}{173.75/69} \quad \# \text{Test stat for 2 vs 3 (SLR vs ANOVA) comparison}
\]

```r
1-pf(0.5024,5,69) \quad \# \text{p-value for 2 vs 3 (SLR vs ANOVA) comparison}
```

\[
\text{[1] 0.7734214}
\]

The Residual SS and Residual DF for each model is from the ANOVA tables provided earlier for each model. We also included a similar test of SLR to the reduced or “null” equal means model.

**QUESTIONS:**

1. Give a conclusion when testing \( H_0 \) : the null model is the true model vs \( H_a \): SLR is the true model. So is increasing model complexity from 1 to 2 parameters appropriate?

2. Give a conclusion when testing \( H_0 \) : the SLR is the true model vs \( H_a \): ANOVA is the true model. So is increasing model complexity from 2 to 7 parameters appropriate?