Glue Strength: AIC Example

We return to the Glue Strength data but this time consider AIC model selection. Recall that the response variable is glue strength. The predictors are glue formulation (A, B, C, D) and application thickness (continuous).

We use both extractAIC() or dredge() from the MuMIn package to calculate AIC. Note that the AIC values do NOT match, but difference (delta) AIC is the same. In other words, the conclusions are the same.

```
library(car)
Glue <- read.csv("C:/hess/STAT512/RNotes/MultReg2/MR2_Glue.csv")
str(Glue)
```

```
## 'data.frame':    20 obs. of  3 variables:
##  $ glue : Factor w/ 4 levels "A","B","C","D": 1 1 1 1 1 2 2 2 2 2 ...
##  $ stren: num  45.5 44.9 48.8 45.1 43.3 48.7 49 50.1 48.5 45.2 ...
##  $ thick: int  13 14 12 12 14 12 10 11 12 14 ...
```

Linear Regression

```
Model1 <- lm(stren ~ thick, data=Glue)
summary(Model1)
```

```
##
## Call:
## lm(formula = stren ~ thick, data = Glue)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.9142 -0.9798  0.2858  0.8334  2.6139
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  61.3267     2.3779  25.790 1.15e-15 ***
## thick        -1.1094     0.1891  -5.867 1.48e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.5 on 18 degrees of freedom
## Multiple R-squared:  0.6566, Adjusted R-squared:  0.6376
## F-statistic: 34.42 on 1 and 18 DF,  p-value: 1.481e-05
```

```
extractAIC(Model1)
```

```
## [1]  2.00000 18.11755
```

```
Anova(Model1, type = 3)
```

```
## Anova Table (Type III tests)
##
## Response: stren
##              Sum Sq Df F value     Pr(>F)
## (Intercept) 1497.01  1 665.134 1.149e-15 ***
```

```
## thick          77.47  1  34.422 1.481e-05 ***
## Residuals      40.51 18
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

ANOVA

```
Model2 <- lm(stren ~ glue, data = Glue)
summary(Model2)
```

```
##
## Call:
## lm(formula = stren ~ glue, data = Glue)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -4.080 -1.260  0.120  1.145  3.920
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   45.520      1.003  45.371   <2e-16 ***
## glueB          2.780      1.419   1.959   0.0677 .
## glueC          3.640      1.419   2.565   0.0207 *
## glueD          1.560      1.419   1.099   0.2878
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.243 on 16 degrees of freedom
## Multiple R-squared:  0.3175, Adjusted R-squared:  0.1895
## F-statistic: 2.481 on 3 and 16 DF,  p-value: 0.09826
```

```
extractAIC(Model2)
```

```
## [1]  4.00000 35.85745
```

```
Anova(Model2, type = 3)
```

```
## Anova Table (Type III tests)
##
## Response: stren
##             Sum Sq Df  F value  Pr(>F)
## (Intercept) 10360.4  1 2058.4844 < 2e-16 ***
## glue          37.5  3    2.4808 0.09826 .
## Residuals     80.5 16
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

ANCOVA NO Interaction

```
Model3 <- lm(stren ~ glue + thick, data = Glue)
summary(Model3)
```

```
## 
## Call:
## lm(formula = stren ~ glue + thick, data = Glue)
## 
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.85815 -0.93808  0.09603  0.78135  2.27007
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  58.6491     2.0793  28.206 2.07e-14 ***
## glueB         1.5681     0.7704   2.036  0.05986 .
## glueC         2.8321     0.7578   3.737  0.00198 **
## glueD         1.3580     0.7483   1.815  0.08960 .
## thick        -1.0099     0.1547  -6.529 9.54e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 1.182 on 15 degrees of freedom
## Multiple R-squared:  0.8223, Adjusted R-squared:  0.775
## F-statistic: 17.36 on 4 and 15 DF,  p-value: 1.688e-05
```

```r
extractAIC(Model3)
```

```
## [1]  5.00000 10.93981
```

```r
Anova(Model3, type = 3)
```

```
## Anova Table (Type III tests)
## 
## Response: stren
##              Sum Sq Df  F value     Pr(>F)
## (Intercept) 1111.80  1 795.5714 2.065e-14 ***
## glue          19.55  3   4.6632   0.01705 *
## thick         59.57  1  42.6236 9.543e-06 ***
## Residuals     20.96 15
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

ANCOVA WITH Interaction

```r
Model4 <- lm(stren ~ glue*thick, data = Glue)
summary(Model4)
```

```
## 
## Call:
## lm(formula = stren ~ glue * thick, data = Glue)
## 
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.84500 -0.75612  0.06574  0.65227  1.85500
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
```

3

```
## (Intercept)   64.0450     7.8833    8.124 3.21e-06 ***
## glueB         -3.1405     9.2522   -0.339   0.7402
## glueC         -6.6046     8.6040   -0.768   0.4576
## glueD         -2.1017     8.4492   -0.249   0.8078
## thick         -1.4250     0.6050   -2.355   0.0363 *
## glueB:thick    0.3568     0.7296    0.489   0.6336
## glueC:thick    0.7463     0.6662    1.120   0.2846
## glueD:thick    0.2638     0.6486    0.407   0.6913
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.21 on 12 degrees of freedom
## Multiple R-squared:  0.8511, Adjusted R-squared:  0.7642
## F-statistic: 9.799 on 7 and 12 DF,  p-value: 0.0003865
```

```r
extractAIC(Model4)
```

```
## [1]  8.00000 13.40662
```

```r
Anova(Model4, type = 3)
```

```
## Anova Table (Type III tests)
##
## Response: stren
##             Sum Sq Df F value   Pr(>F)
## (Intercept) 96.626  1 66.0022 3.21e-06 ***
## glue         1.785  3  0.4065  0.75115
## thick        8.122  1  5.5482  0.03635 *
## glue:thick   3.395  3  0.7729  0.53116
## Residuals   17.568 12
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

AIC Comparison with MuMIn package

To use dredge(), you supply the "full" model (the "largest" or most complicated model you want to consider). MuMIn handles factors and produces a nice summary table. For continuous predictors, the partial regression coefficients are shown. For categorical predictors (factors), the + in the summary table just indicates that predictor is included in the model. By default, dredge() will rank models by AICc. That choice is completely reasonable, but I use AIC here to compare to other methods. Note that variance terms is included in the count of parameters.

```r
library(MuMIn)
options(na.action = "na.fail")
dredge(Model4, rank="AIC")
```

```
## Fixed term is "(Intercept)"

## Global model call: lm(formula = stren ~ glue * thick, data = Glue)
## ---
## Model selection table
##   (Int) glu    thc glu:thc df  logLik  AIC delta weight
## 4 58.65   + -1.010         6 -28.849 69.7  0.00  0.758
## 8 64.04   + -1.425       + 9 -27.082 72.2  2.47  0.221
## 3 61.33     -1.109         3 -35.438 76.9  7.18  0.021
```

4

```
## 2 45.52   +                  5 -42.307 94.6 24.92  0.000
## 1 47.52                       2 -46.127 96.3 26.56  0.000
## Models ranked by AIC(x)
```