Glue Strength Example: ANCOVA with 4 groups

The response variable is glue strength. The predictors are glue formulation (A, B, C, D) and application thickness (continuous). The research question is focused on differences between the glues but thickness is included in the model as a covariate.

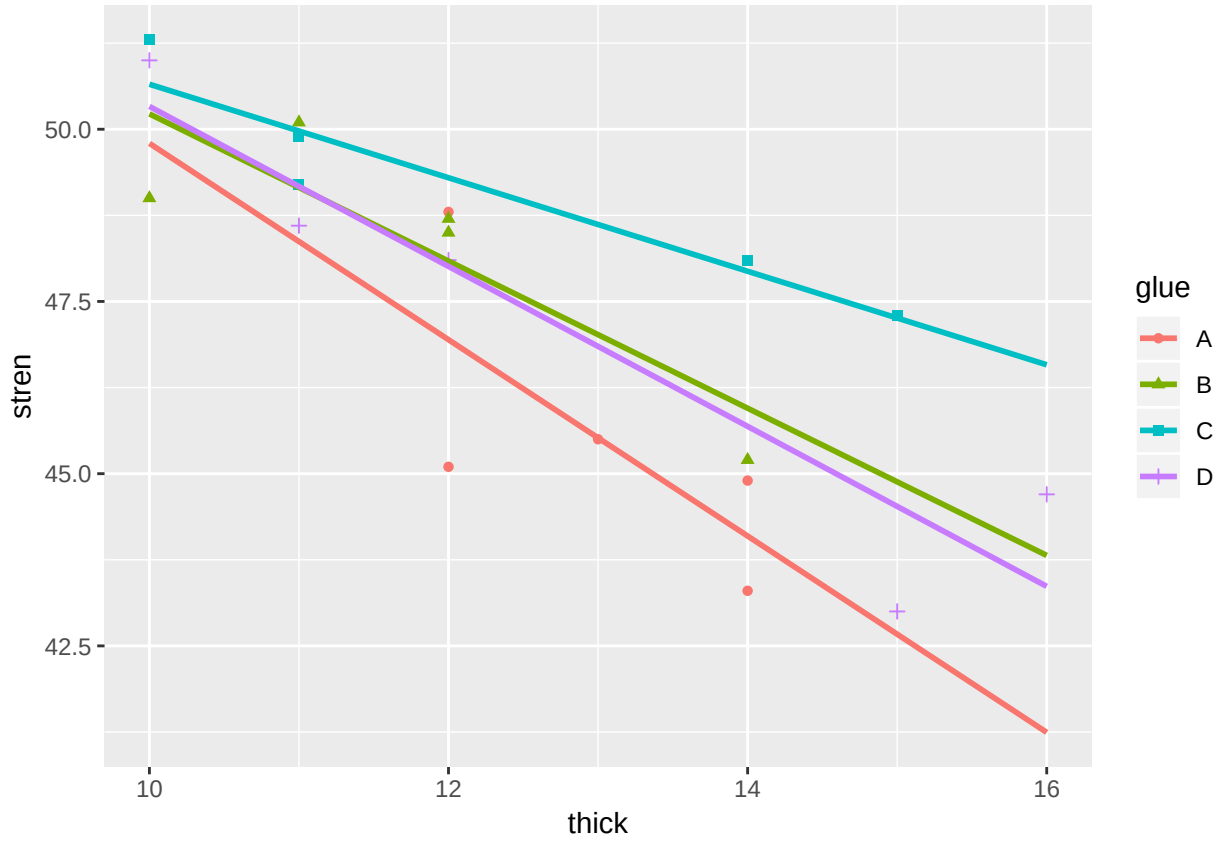
Based on the research goal (and the fact that the glue:thick interaction is not significant), we are probably most interested in Model2 (no interaction) and the emmeans comparisons.

However, we use this data to illustrate running tests of slopes based on Models 3 and 4 (with interaction).

```
library(ggplot2)
library(dplyr)
library(car)
library(emmeans)
library(multcompView)
Glue <- read.csv("~/Dropbox/STAT512/Lectures/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 ...
```

```
p <- qplot(thick, stren, shape = glue, color = glue, data = Glue)
p + geom_smooth(method = "lm", se = FALSE, fullrange = T)
```



```r
#Summary statistics using dplyr
SumStats <- summarise(group_by(Glue, glue),
                      n = n(),
                      mean = mean(stren),
                      sd = sd(stren),
                      SE = sd/sqrt(n))
SumStats
```

```
## # A tibble: 4 x 5
##   glue      n  mean    sd    SE
##   <fct> <int> <dbl> <dbl> <dbl>
## 1 A         5  45.5  2.02 0.901
## 2 B         5  48.3  1.84 0.823
## 3 C         5  49.2  1.56 0.697
## 4 D         5  47.1  3.20 1.43
```

Model1: One-way ANOVA

```r
model1 <- lm(stren ~ glue, data = Glue)
Anova(model1, 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
```

```r
emmeans(model1, pairwise ~ glue)
```

```
## $emmeans
##  glue emmean       SE df lower.CL upper.CL
##  A     45.52 1.003295 16 43.39311 47.64689
##  B     48.30 1.003295 16 46.17311 50.42689
##  C     49.16 1.003295 16 47.03311 51.28689
##  D     47.08 1.003295 16 44.95311 49.20689
##
## Confidence level used: 0.95
##
## $contrasts
##  contrast estimate       SE df t.ratio p.value
##  A - B       -2.78 1.418873 16  -1.959  0.2438
##  A - C       -3.64 1.418873 16  -2.565  0.0871
##  A - D       -1.56 1.418873 16  -1.099  0.6949
##  B - C       -0.86 1.418873 16  -0.606  0.9286
##  B - D        1.22 1.418873 16   0.860  0.8251
##  C - D        2.08 1.418873 16   1.466  0.4794
##
## P value adjustment: tukey method for comparing a family of 4 estimates
```

```
extractAIC(model1)
```

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

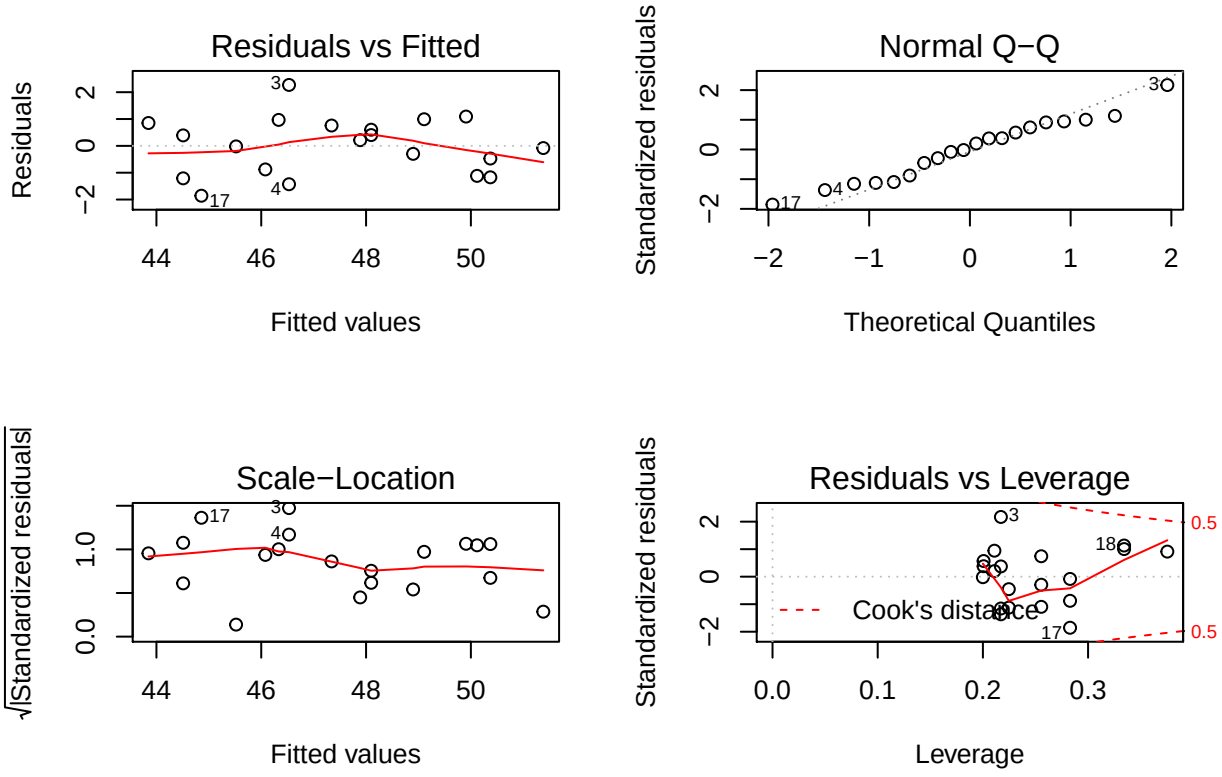
Model2: ANCOVA NO Interaction

Note that after accounting for thick, we are able to detect significant differences between the glues. The no interaction model forces the slope to be the same for all groups.

```
model2 <- lm(stren ~ glue + thick, data = Glue)
Anova(model2, 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
```

```
par(mfrow = c(2,2))
plot(model2)
```




```
emout <- emmeans(model2, pairwise ~ glue)
emout
```

```
## $emmeans
##  glue   emmean        SE df lower.CL upper.CL
##  A    46.07546 0.5354766 15 44.93412 47.21680
##  B    47.64354 0.5381512 15 46.49650 48.79059
##  C    48.90752 0.5300869 15 47.77766 50.03737
##  D    47.43348 0.5314395 15 46.30074 48.56621
##
## Confidence level used: 0.95
##
## $contrasts
##  contrast   estimate        SE df t.ratio p.value
##  A - B    -1.5680822 0.7703581 15  -2.036  0.2189
##  A - C    -2.8320548 0.7578311 15  -3.737  0.0095
##  A - D    -1.3580137 0.7482982 15  -1.815  0.3048
##  B - C    -1.2639726 0.7502145 15  -1.685  0.3650
##  B - D     0.2100685 0.7634937 15   0.275  0.9924
##  C - D     1.4740411 0.7533974 15   1.957  0.2473
##
## P value adjustment: tukey method for comparing a family of 4 estimates
```

```
CLD(emout$emmeans)
```

```
##  glue   emmean        SE df lower.CL upper.CL .group
##  A    46.07546 0.5354766 15 44.93412 47.21680  1
##  D    47.43348 0.5314395 15 46.30074 48.56621  12
##  B    47.64354 0.5381512 15 46.49650 48.79059  12
##  C    48.90752 0.5300869 15 47.77766 50.03737   2
##
## Confidence level used: 0.95
## P value adjustment: tukey method for comparing a family of 4 estimates
## significance level used: alpha = 0.05
```

```
extractAIC(model2)
```

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

Model3: ANCOVA WITH Interaction #1

The model with interaction allows for different slopes and intercepts for each group. We can get estimates and pairwise comparisons of slopes using the emtrends function. Given that thickness is not of primary research interest, inference about slopes would probably not be of interest here.

```
model3 <- lm(stren ~ glue*thick, data = Glue)
Anova(model3, 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
```

```
extractAIC(model3)
```

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

```
summary(model3)
```

```
##
## 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|)
## (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
```

```
emtrends(model3, pairwise ~ glue, var = "thick")
```

```
## $emtrends
##  glue thick.trend        SE df  lower.CL   upper.CL
##  A     -1.4250000 0.6049746 12 -2.743126 -0.1068737
##  B     -1.0681818 0.4078738 12 -1.956862 -0.1795012
##  C     -0.6787234 0.2790540 12 -1.286730 -0.0707171
##  D     -1.1611940 0.2337221 12 -1.670431 -0.6519574
##
## Confidence level used: 0.95
##
## $contrasts
##  contrast   estimate        SE df t.ratio p.value
##  A - B    -0.3568182 0.7296268 12  -0.489  0.9601
##  A - C    -0.7462766 0.6662322 12  -1.120  0.6847
##  A - D    -0.2638060 0.6485524 12  -0.407  0.9762
##  B - C    -0.3894584 0.4941985 12  -0.788  0.8585
##  B - D     0.0930122 0.4700926 12   0.198  0.9971
##  C - D     0.4824706 0.3640015 12   1.325  0.5653
##
## P value adjustment: tukey method for comparing a family of 4 estimates
```

Model4: ANCOVA WITH Interaction #2 (Alternate Parameterization)

This model is helpful for getting the slope and intercept estimates directly. We can also use this model to run pairwise comparisons of slopes (and intercepts). Primarily for illustration for this data, but may be of interest in other scenarios.

```
model4 <- lm(stren ~ -1 + glue + thick:glue, data = Glue)
extractAIC(model4)
```

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

```
summary(model4)
```

```
##
## Call:
## lm(formula = stren ~ -1 + glue + thick:glue, 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|)
## glueA        64.0450     7.8833   8.124 3.21e-06 ***
## glueB        60.9045     4.8432  12.575 2.87e-08 ***
## glueC        57.4404     3.4472  16.663 1.16e-09 ***
## glueD        61.9433     3.0402  20.375 1.12e-10 ***
## glueA:thick  -1.4250     0.6050  -2.355 0.036349 *
## glueB:thick  -1.0682     0.4079  -2.619 0.022430 *
## glueC:thick  -0.6787     0.2791  -2.432 0.031603 *
## glueD:thick  -1.1612     0.2337  -4.968 0.000326 ***
## ---
## 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:  0.9996, Adjusted R-squared:  0.9994
## F-statistic:  3864 on 8 and 12 DF,  p-value: < 2.2e-16
```

```
#Example of Pairwise comparisons of interecepts and slopes
#Primarily for illustration probably not of interest here.
BintvAint <- c(-1, 1, 0, 0, 0, 0, 0, 0)
lht(model4, BintvAint, rhs = c(0))
```

```
## Linear hypothesis test
##
## Hypothesis:
## - glueA  + glueB = 0
##
## Model 1: restricted model
## Model 2: stren ~ -1 + glue + thick:glue
##
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1     13 17.736
## 2     12 17.568  1   0.16867 0.1152 0.7402
```

```r
BslopevAslope <- c(0, 0, 0, 0, -1, 1, 0, 0)
lht(model4, BslopevAslope, rhs = c(0))
```

```
## Linear hypothesis test
##
## Hypothesis:
## - glueA:thick  + glueB:thick = 0
##
## Model 1: restricted model
## Model 2: stren ~ -1 + glue + thick:glue
##
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1     13 17.918
## 2     12 17.568  1   0.35013 0.2392 0.6336
```