MAT325 Project 6

Jordan Badstuebner

22 August 2020

```
#############
####  1  ####
#############
```

```
##
## Call:
## lm(formula = df1_g$S_Uniformity ~ df1_g$A + df1_g$B + df1_g$AB)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -5.000 -1.375 -0.250  1.812  4.000
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  37.0000     0.6353  58.237 4.34e-16 ***
## df1_g$A       1.3750     0.6353   2.164 0.051322 .
## df1_g$B       3.0000     0.6353   4.722 0.000495 ***
## df1_g$AB      8.3750     0.6353  13.182 1.69e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.541 on 12 degrees of freedom
## Multiple R-squared:  0.9436, Adjusted R-squared:  0.9295
## F-statistic: 66.92 on 3 and 12 DF,  p-value: 9.213e-08
```

We have $p - val = 1.049 \times 10^{-5}$

We proceed to create an ANOVA table for this experiment to test the main and interaction effects for significace. Use $\alpha = 0.01$.

Step 1:

H_o : The main and interaction effects are not significant.

H_a : The main and interaction effects are significant.

Step 2:

```
## Analysis of Variance Table
##
## Response: df1_g$S_Uniformity
##           Df  Sum Sq Mean Sq  F value    Pr(>F)
```

```
## df1_g$A    1   30.25   30.25   6.7222 0.0319774 *
## df1_g$B    1  144.00  144.00  32.0000 0.0004776 ***
## df1_g$C    1   12.25   12.25   2.7222 0.1375685
## df1_g$AB   1 1122.25 1122.25 249.3889 2.584e-07 ***
## df1_g$AC   1    1.00    1.00   0.2222 0.6499402
## df1_g$BC   1   12.25   12.25   2.7222 0.1375685
## df1_g$ABC  1   16.00   16.00   3.5556 0.0960716 .
## Residuals  8   36.00    4.50
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Step 3: Observe the p and f values shown in the ANOVA results above.

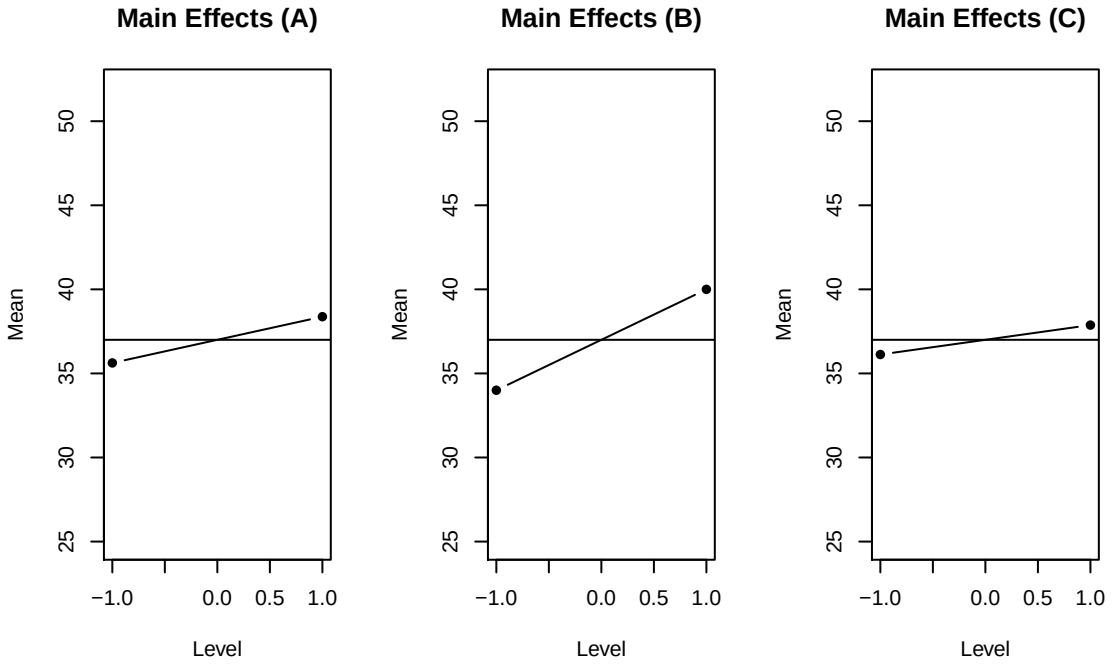
Step 4: We reject the H_o for any $p - val < \alpha = 0.01$.

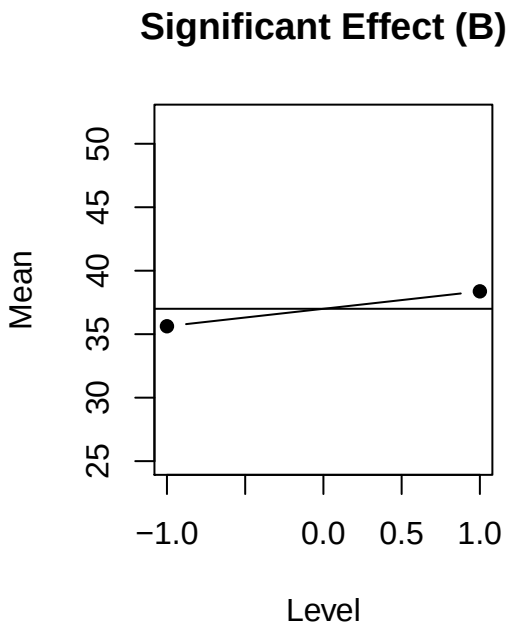
Step 5: There is sufficient evidence to suggest at a significance level of 0.01 that the main effect B and the interaction effect AB are significant.

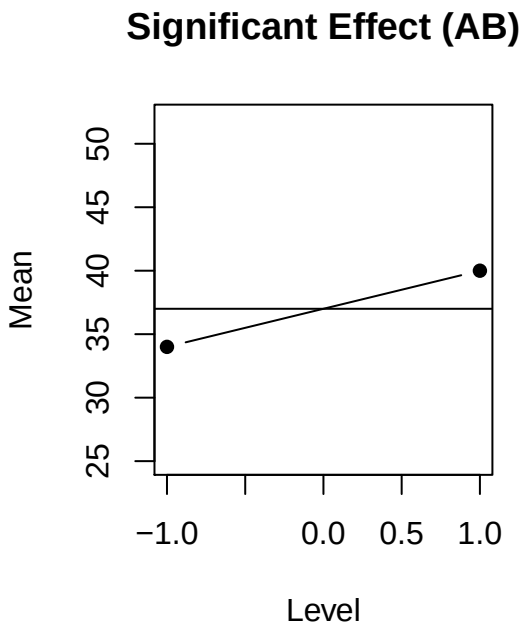
Let us use only the reduced model.

```
## Analysis of Variance Table
##
## Response: df1_g$S_Uniformity
##           Df  Sum Sq Mean Sq  F value    Pr(>F)
## df1_g$A    1   30.25   30.25   4.6839 0.0513222 .
## df1_g$B    1  144.00  144.00  22.2968 0.0004953 ***
## df1_g$AB   1 1122.25 1122.25 173.7677 1.686e-08 ***
## Residuals 12   77.50    6.46
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

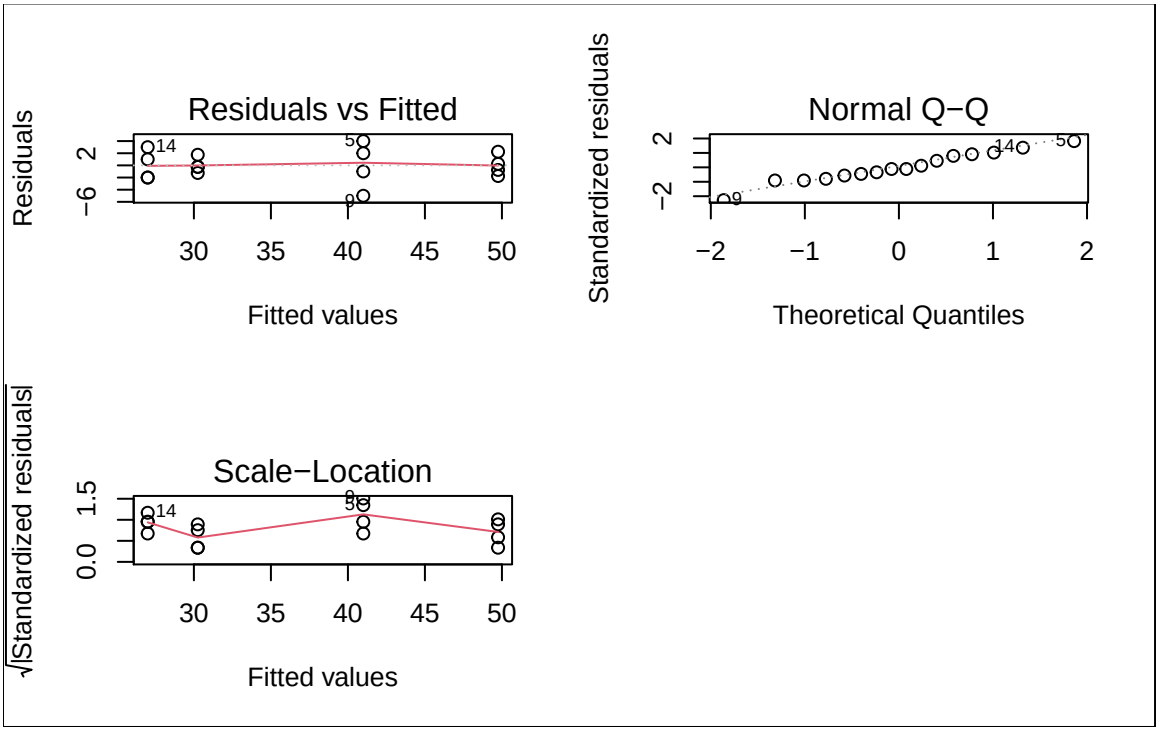
The plots of the main effects and significant effects are shown below:







The Residuals plots of the model are shown below:



There is nothing unusual about the plots. All assumptions appear to be met.

Finally, a Shapiro-wilks normality test is run below:

```
##
##  Shapiro-Wilk normality test
```

```
##
## data:  residuals(lm1_r)
## W = 0.97406, p-value = 0.8994
```

Since we have $p - val = 0.8994 > \alpha = 0.05$ we cannot conclude that the sample deviates from normality.

```
##     X  A  B  C AB AC BC ABC  EC50
## 1 (1) -1 -1 -1  1  1  1  -1 23100
## 2   a  1 -1 -1 -1 -1  1   1 43000
## 3   b -1  1 -1 -1  1 -1   1 71400
## 4  ab  1  1 -1  1 -1 -1  -1 76000
## 5   c -1 -1  1  1 -1 -1   1 37000
## 6  ac  1 -1  1 -1  1 -1  -1 33200
```

```
##
## Call:
## lm(formula = df2$EC50 ~ df2$A + df2$B + df2$C + df2$AB + df2$AC +
##     df2$BC + df2$ABC)
##
## Residuals:
## ALL 8 residuals are 0: no residual degrees of freedom!
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)    39650         NA      NA       NA
## df2$A           2525         NA      NA       NA
## df2$B           5575         NA      NA       NA
## df2$C         -13725         NA      NA       NA
## df2$AB         -1500         NA      NA       NA
## df2$AC         -3600         NA      NA       NA
## df2$BC        -14750         NA      NA       NA
## df2$ABC         2325         NA      NA       NA
##
## Residual standard error: NaN on 0 degrees of freedom
## Multiple R-squared:      1,  Adjusted R-squared:    NaN
## F-statistic:   NaN on 7 and 0 DF,  p-value: NA
```

```
##      A      B     AB      C     AC     BC    ABC
##   5050  11150  -3000 -27450  -7200 -29500   4650
## attr(,"mean")
##
## 39650
```

```
##
## Call:
## lm(formula = df2$EC50 ~ df2$A + df2$B + df2$C + df2$AB + df2$AC)
##
## Residuals:
##      1      2      3      4      5      6      7      8
## -17075 -12425  17075  12425  17075  12425 -17075 -12425
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)    39650      10559   3.755   0.0642 .
## df2$A           2525      10559   0.239   0.8333
## df2$B           5575      10559   0.528   0.6502
## df2$C         -13725      10559  -1.300   0.3233
## df2$AB         -1500      10559  -0.142   0.9000
## df2$AC         -3600      10559  -0.341   0.7656
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 29860 on 2 degrees of freedom
## Multiple R-squared:  0.5195, Adjusted R-squared:  -0.6818
## F-statistic: 0.4324 on 5 and 2 DF,  p-value: 0.8055
```

```
## Analysis of Variance Table
##
## Response: df2$EC50
##           Df     Sum Sq    Mean Sq F value Pr(>F)
## df2$A      1   51005000   51005000  0.0572 0.8333
## df2$B      1  248645000  248645000  0.2788 0.6502
## df2$C      1 1507005000 1507005000  1.6897 0.3233
## df2$AB     1   18000000   18000000  0.0202 0.9000
## df2$AC     1  103680000  103680000  0.1162 0.7656
## Residuals  2 1783745000  891872500
```