Battery Operating Conditions Example

May 11, 2020

```
In [14]: library(MASS)
```

0.1 Factorial Models

0.1.1 Battery Operating Conditions Example

Engineer is investigating **two factors** that may contribute towards battery life:

1. material from which the battery is made
2. temperature in the end-use environment

There are three different material types (M_1, M_2, M_3) and three temperatures (of particular interest (T_1, T_2, T_3).

Let's call **material** Factor A. Therefore $a = 3$

Let's call **temperature** Factor B. Therefore $b = 3$

Entering the data into R:

```
In [1]: battery.life <- c(130, 155, 74, 180, 150, 188, 159, 126,
                          138, 110, 168, 160, 34, 40, 80, 75, 136,
                          122, 106, 115, 174, 120, 150, 139,
                          20, 70, 82, 58, 25, 70, 58, 45, 96, 104, 82, 60)
```

```
In [2]: length(battery.life)
```

36

The data entered above was put in **column order**

So we have 36 observations in total.

$abn = 36, ab = 9, n = 4$

There are four observations per i, j cell.

0.2 Coding in the factors for material and temperature:

Since the data has been entered column by column, you will have 12 values for each **temperature**, going 1, 2, 3.

The material type will be 4 values for M1, 4 values for M2, 4 values for M3, then back to 4 values for M1 again.

```
In [3]: temp <- factor(c(rep(1,12),rep(2,12),rep(3,12)))
```

```
In [4]: material <- factor(c( rep(rep(1:3, rep(4,3)),3) ))
```

```
In [5]: battery <- data.frame(battery.life, temp, material)
```

```
In [6]: # cleaning up
        rm(battery.life, temp, material)
        attach(battery)
```

0.3 Running ANOVA with and without the interaction terms

```
In [10]: battery.int.aov <- aov(battery.life ~ temp + material + temp:material)

         battery.no.int.aov <- aov(battery.life ~ temp + material)
```

0.4 First looking at the ANOVA with the interaction terms

Recall the following **three** hypotheses:

1. Factor A (**material**), $H_{0A} : \tau_1 = \tau_2 = \tau_3 = 0$
2. Factor B (**temperature**), $H_{0B} : \beta_1 = \beta_2 = \beta_3 = 0$
3. Factor AB (**interaction**), $H_{0AB} : (\tau\beta)_{ij} = 0$, for $i = 1, 2, 3$, and $j = 1, 2, 3$

```
In [11]: summary(battery.int.aov)
```

```
              Df Sum Sq Mean Sq F value   Pr(>F)
temp           2  39119   19559  28.968 1.91e-07 ***
material       2  10684    5342   7.911  0.00198 **
temp:material  4   9614    2403   3.560  0.01861 *
Residuals     27  18231     675
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

0.5 Summary of this ANOVA table

Looking first at the F-stat for the first hypothesis test, H_0A, where we look at **material**. The number of degrees of freedom is $a - 1 = 3 - 1 = 2$. The F-value, under the null hypothesis is drawn from the $F_{2,27}$ distribution, whereby we can reject the null hypothesis at the 1% level of significance, where the null hypothesis stated that the treatment means within the material factor were all the same, and thus that the treatment effects for each level of the factor (i.e. each material) was zero. We accept the alternative hypothesis that at least one of the material treatment effects is non-zero.

Secondly, looking at the F-stat for the second hypothesis test, H_{0B}, where we're looking at **temp**. The number of degrees of freedom associated with the temp sum of squares is $b - 1 = 3 - 1 = 2$. The F-score associated with temp under the null hypothesis is drawn from the $F_{2,27}$ distribution. The p-value associated with temp is highly significant, and therefore we can reject the H_{0B} null hypothesis at the 0.1% level of significance. The null hypothesis stated that there was no difference in the treatment means between temperature treatment groups, and therefore that the temperature treatment effects were zero for all levels within the group (i.e. for all three temperatures). We therefore accept the alternative hypothesis that at least one of the treatment effects for a temperature level is non-zero.

Lastly, looking at the third hypothesis, H_{0AB}, where we're looking at the **interaction** term. The number of degrees of freedom associated with the interaction term is $(ab - 1) - (a - 1) - (b - 1) = (a - 1)(b - 1) = 2 \times 2 = 4$. The F-score under the null hypothesis is drawn from the $F_{4,27}$ distribution, where the null hypothesis is that the interaction treatment effects are zero for all 9 interations. The p-value associated with this F-test means we can reject the null hypothesis at the 5% significance level, and therefore accept the alternative hypothesis that at least one of the interaction treatment effects is non-zero.

0.6 Looking at the table of means.

```
In [13]: model.tables(battery.int.aov, type='means')
```

```
Tables of means
Grand mean

105.5278

 temp
temp
     1      2      3
144.83 107.58  64.17

 material
material
     1      2      3
 83.17 108.33 125.08

 temp:material
    material
temp 1      2      3
   1 134.75 155.75 144.00
   2  57.25 119.75 145.75
   3  57.50  49.50  85.50
```

0.7 Two interesting types of plot for inspecting two-factor ANOVA

1. `plot.design`

2. `interaction.plot`

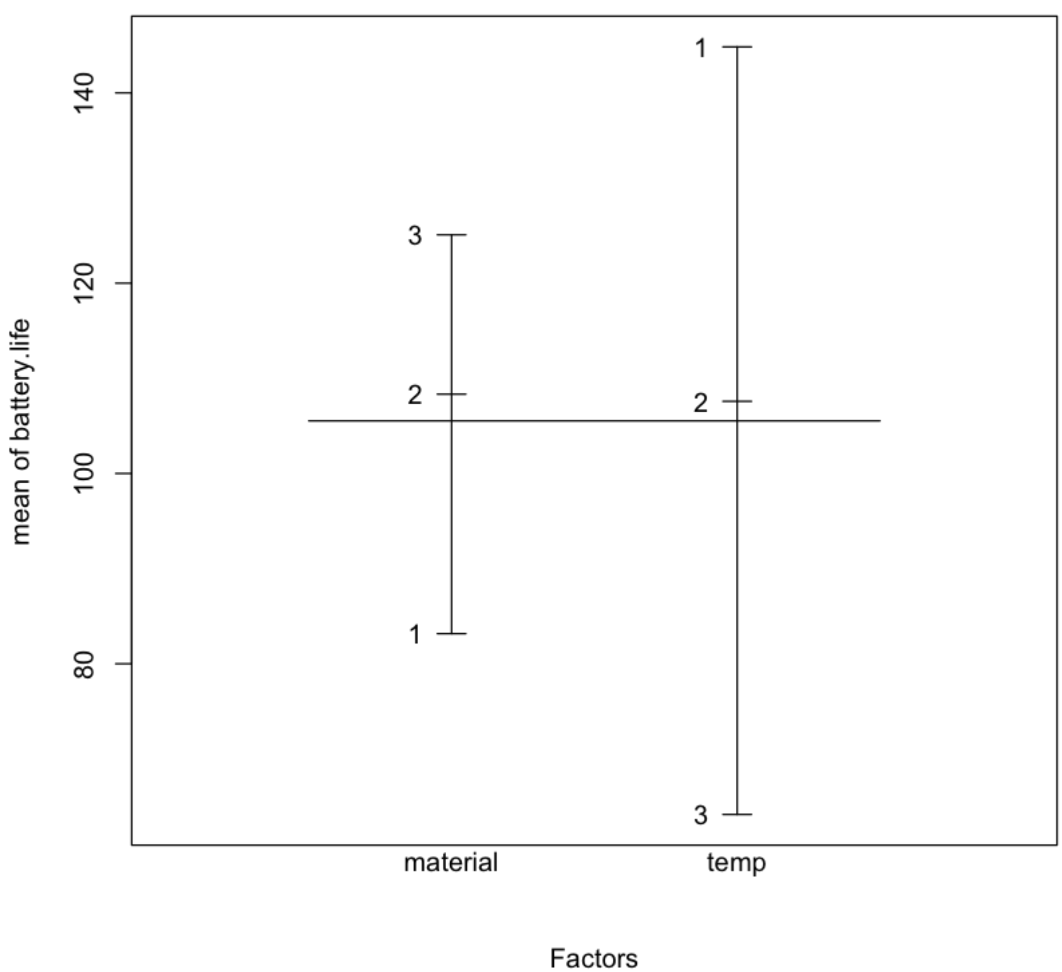
The `plot.design` shows the overall mean of the dependent variable, over the top of the treatment means for each level of both treatments / factors.

Can clearly see variability between the two factors, whereby we can see that temperature 1 provides a higher battery life than temperature 3.

Can also clearly see that material 3 provides the highest battery life too.

This plot doesn't show the interaction effects.

```
In [18]: plot.design(battery.life ~ material + temp)
```

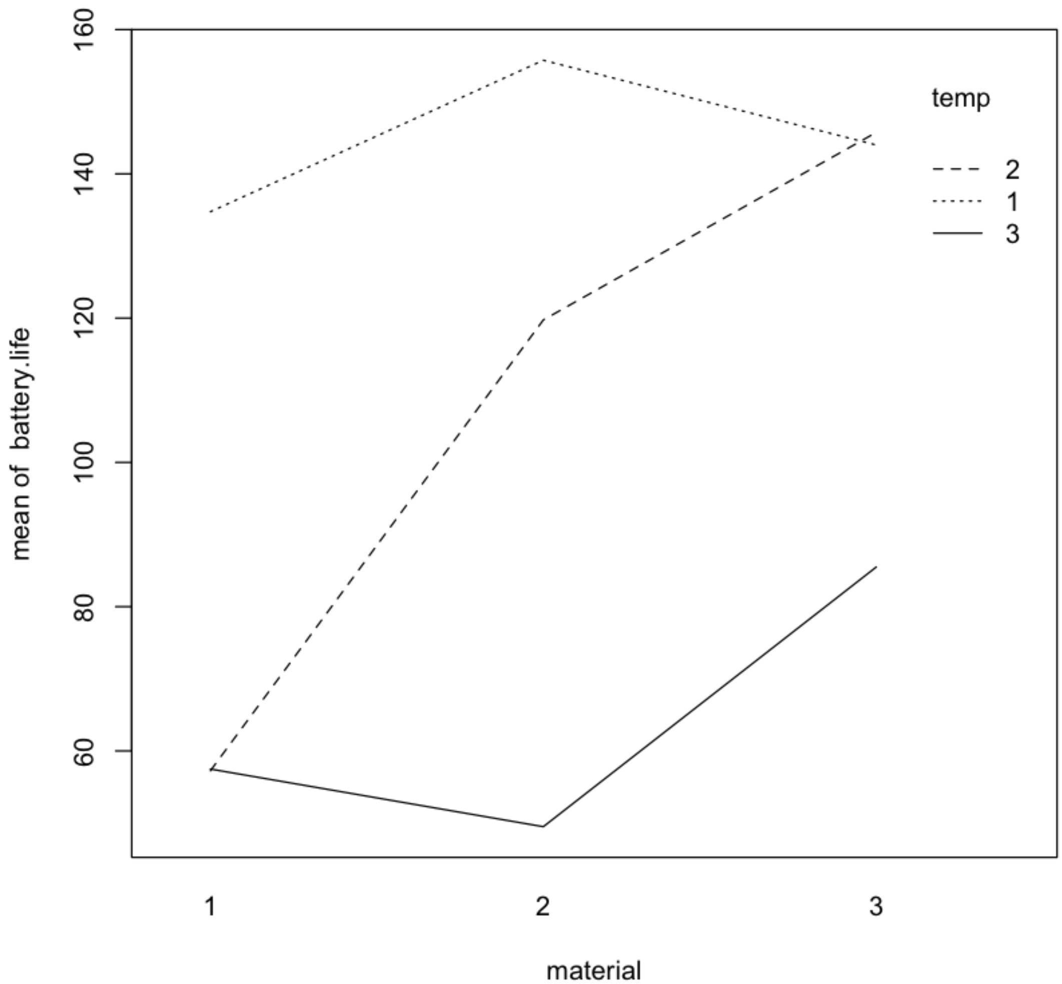


The `interaction.plot` plots (both ways round), shows that at temperature 2, changing material has a more than doubles the battery life.

At temperature 1, there's not much variation between the materials.

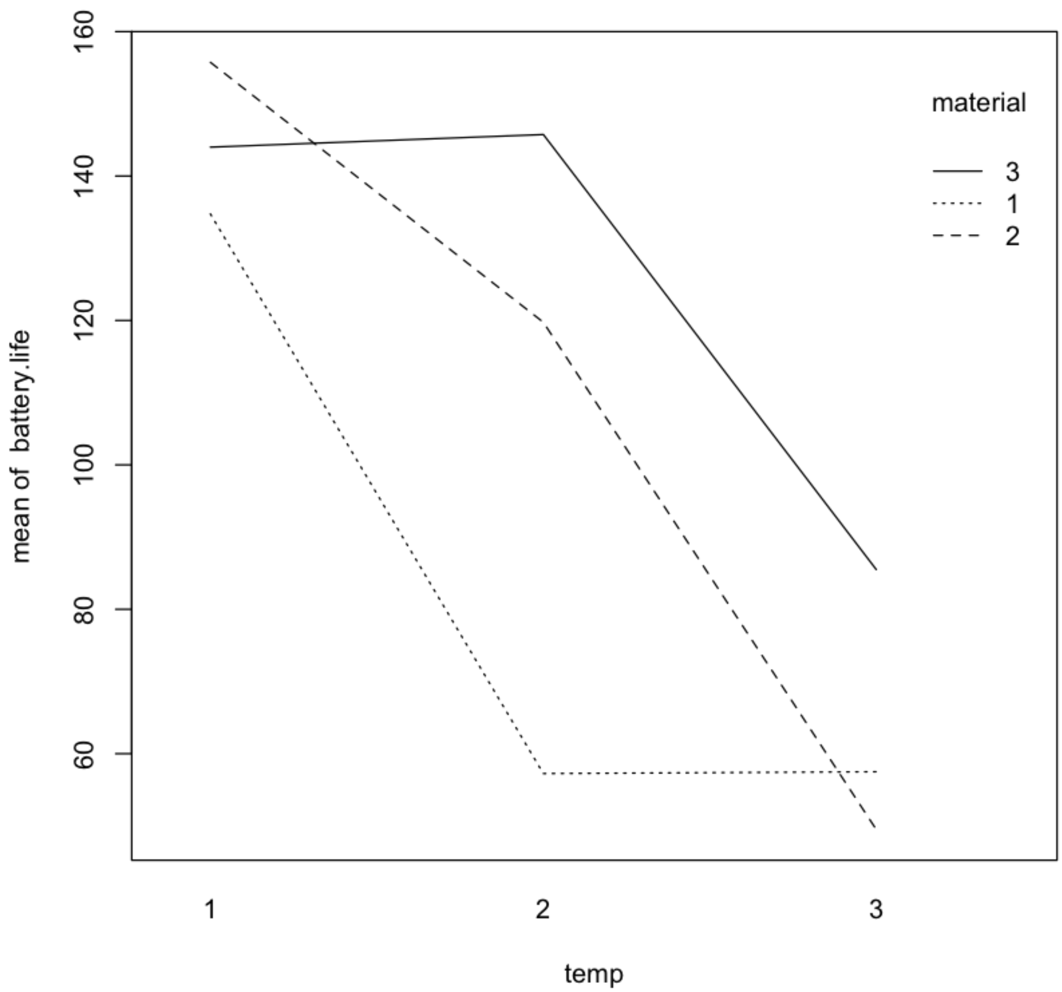
At temperature 3, material 3 seems to perform better than 1 and 2.

```
In [16]: interaction.plot(material, temp, battery.life)
```



There's a very clear trend that as the temperature goes from 1 -> 2 -> 3, (i.e. as temperature increases) that battery life decreases accross all materials.

```
In [17]: interaction.plot(temp, material, battery.life)
```



0.8 Looking at the residuals

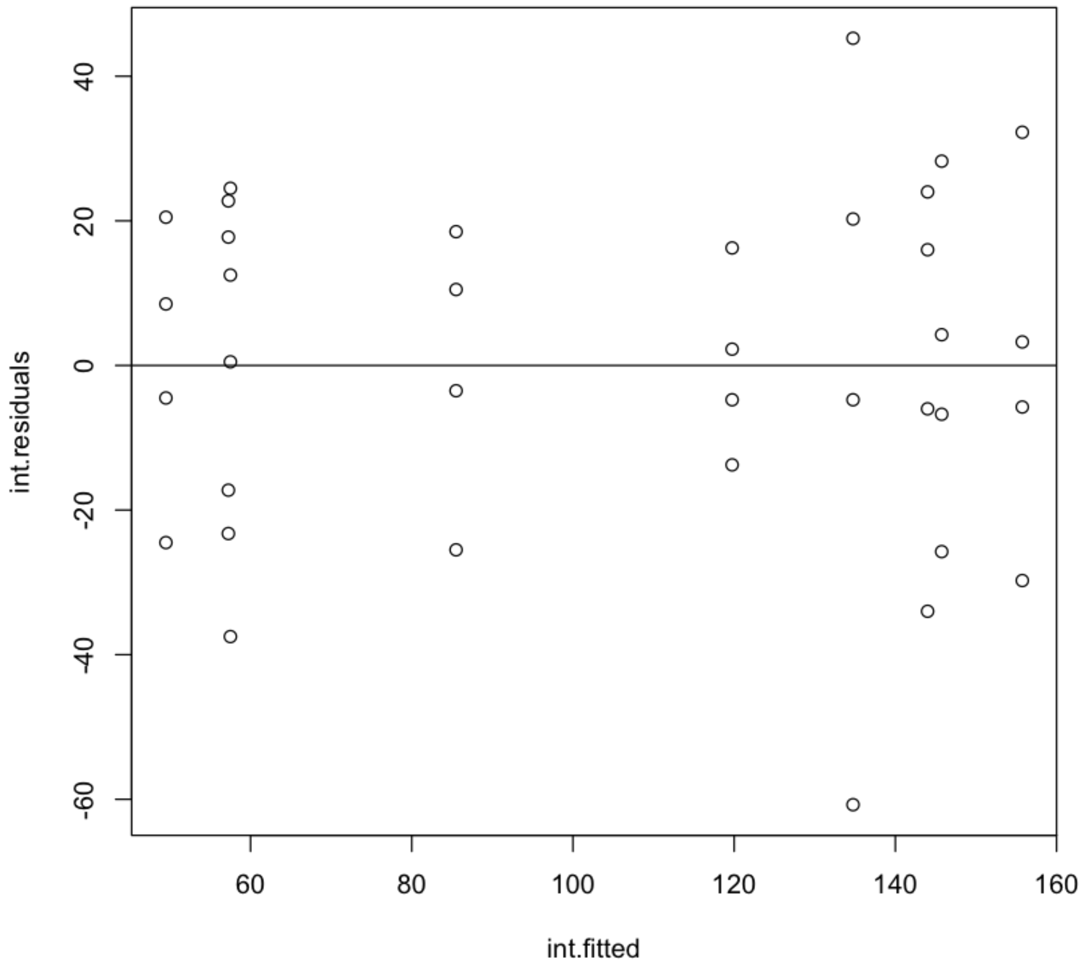
You can get the residuals directly from the AOV object

You can also get the fitted values directly from the AOV obect

```
In [20]: int.residuals <- residuals(battery.int.aov)

         int.fitted <- fitted(battery.int.aov)
```

0.9 Plotting the residuals against the fitted values

```
In [24]: plot(int.fitted, int.residuals)
         abline(h=0)
```

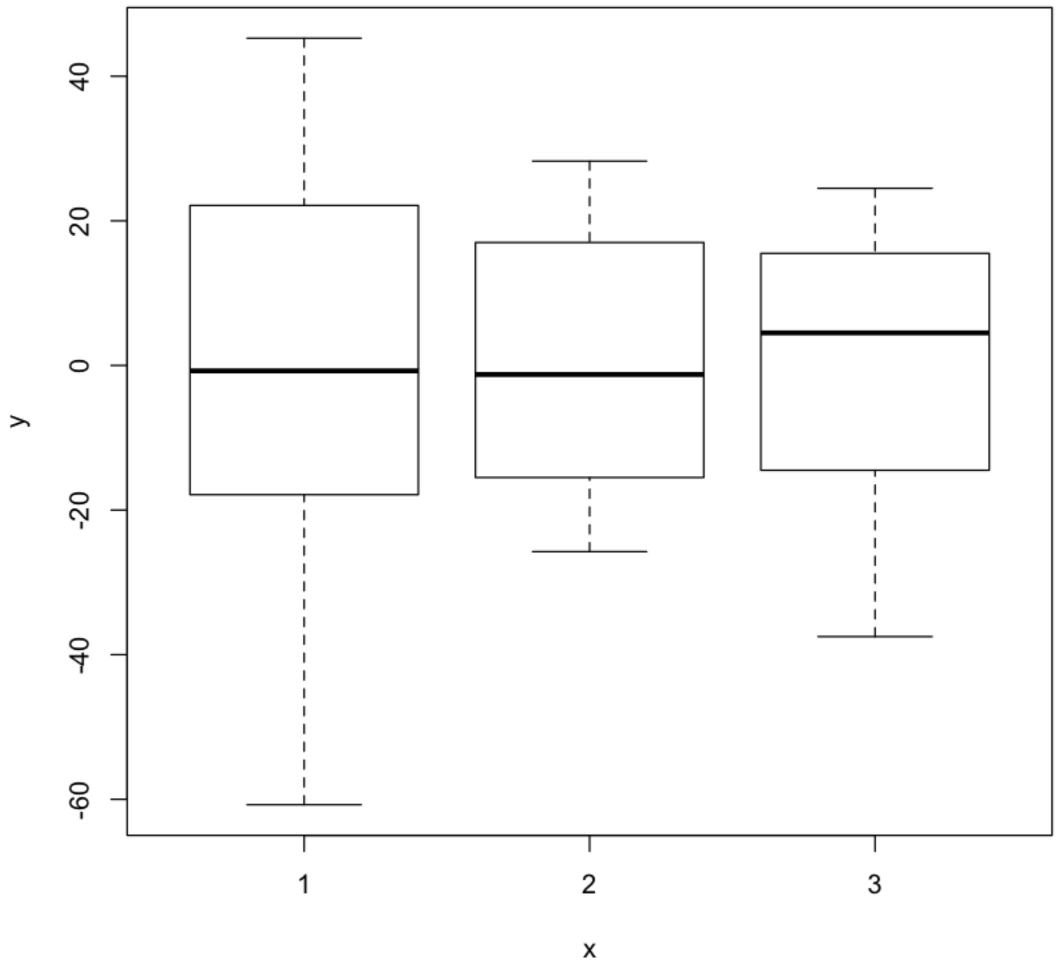


To inspect the assumption that the error terms are mean zero with constant variance, we inspect the residuals plotted against the fitted values, as a test of model adequecy (i.e are the assumptions behind the model holding up).

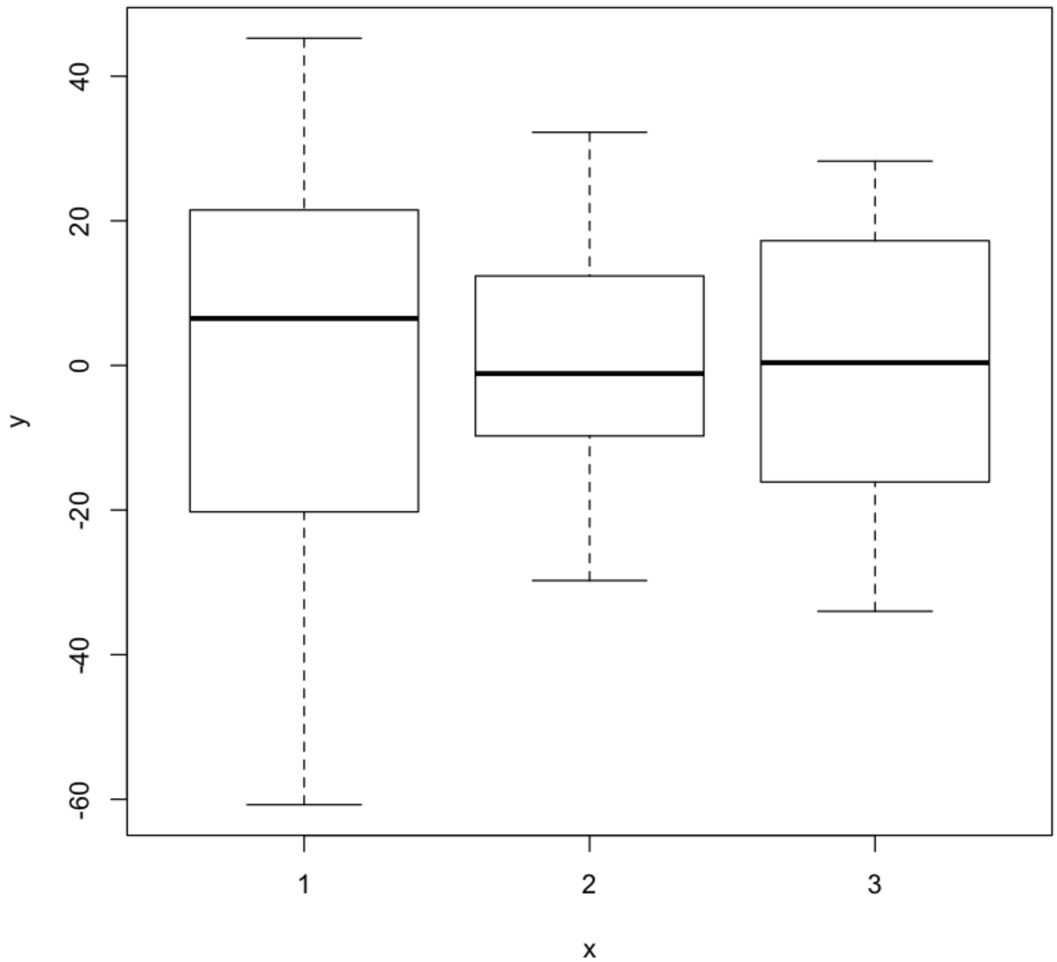
You can clearly see even scatter around zero, where as the fitted values increase in value, the variance appears to be fairly consistent.

0.10 And now plotting the residuals against the factors

```
In [25]: plot(temp, int.residuals)
```




```
In [26]: plot(material, int.residuals)
```



By looking at the plots of the residuals against the factors, we see slighly more variation in the residuals for both the first temperature (the lowest temperature) and the first material (M_1).

0.11 The no interaction model

Note, here's how the model with interactions and without interactions look:

With interactions (i.e. including $(\tau\beta)_{ij}$):

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk},$$

where $i = 1, 2, ..., a$, $j = 1, 2, ..., b$, and $k = 1, 2, ..., n$.

Without interactions:

$$y_{ijk} = \mu + \tau_i + \beta_j + \epsilon_{ijk}$$

Where parameter estimates for μ, τ_i, β_j will be the same, since the interaction term disappears due to sum to zero constraints when finding those estimates.

The one estimate that will change will be $\hat{y}_{ijk}$.

$$\hat{y}_{ijk} = \hat{\mu} + \hat{\tau}_i + \hat{\beta}_j,$$

where:

$$\hat{\mu} = \bar{y}_{...},$$

$$\hat{\tau}_i = \bar{y}_{i..} - \bar{y}_{...},$$

$$\hat{\beta}_j = \bar{y}_{.j.} - \bar{y}_{...},$$

therefore:

$$\hat{y}_{ijk} = \bar{y}_{...} + \bar{y}_{i..} - \bar{y}_{...} + \bar{y}_{.j.} - \bar{y}_{...}$$

$$\hat{y}_{ijk} = \bar{y}_{i..} + \bar{y}_{.j.} - \bar{y}_{...},$$

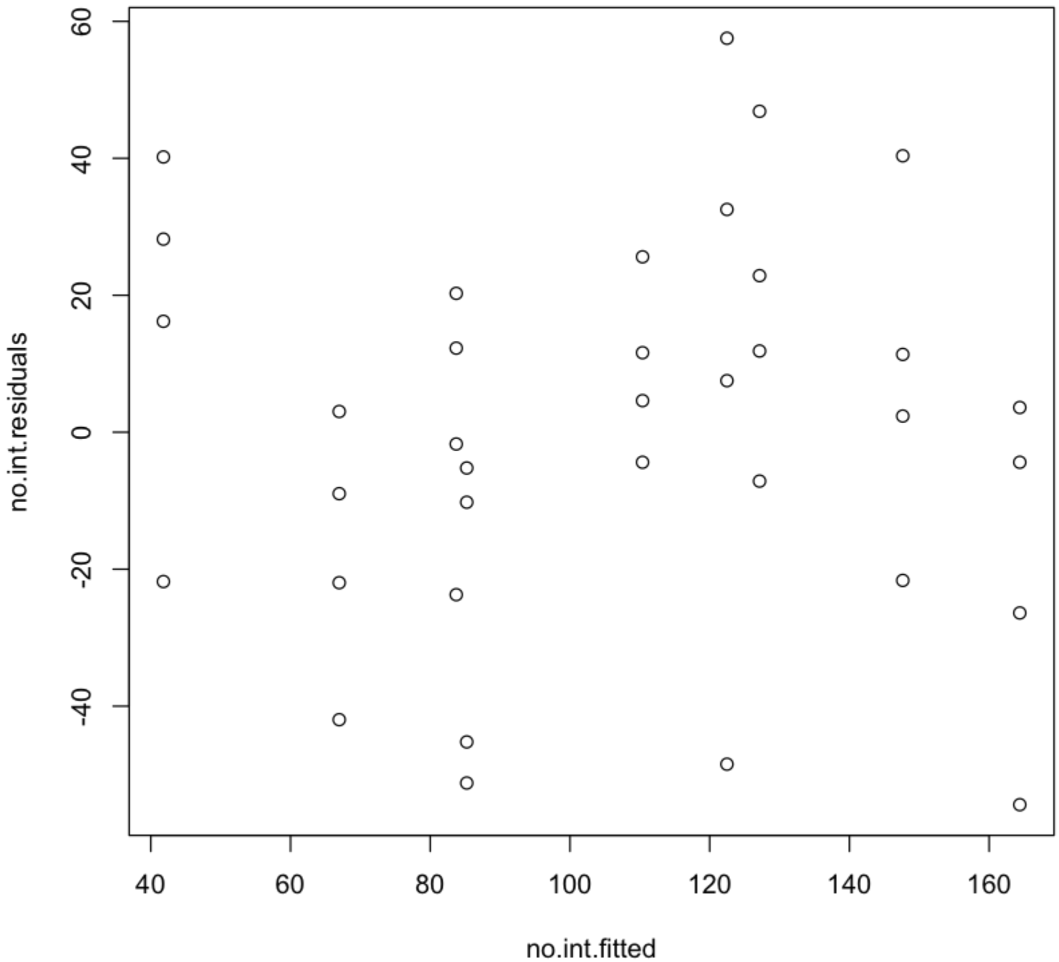
$i = 1, 2, ..., a,\ j = 1, 2, ..., b,\ k = 1, 2, ..., n.$

```
In [27]: summary(battery.no.int.aov)
```

```
            Df Sum Sq Mean Sq F value   Pr(>F)
temp         2  39119   19559  21.776 1.24e-06 ***
material     2  10684    5342   5.947  0.00651 **
Residuals   31  27845     898
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

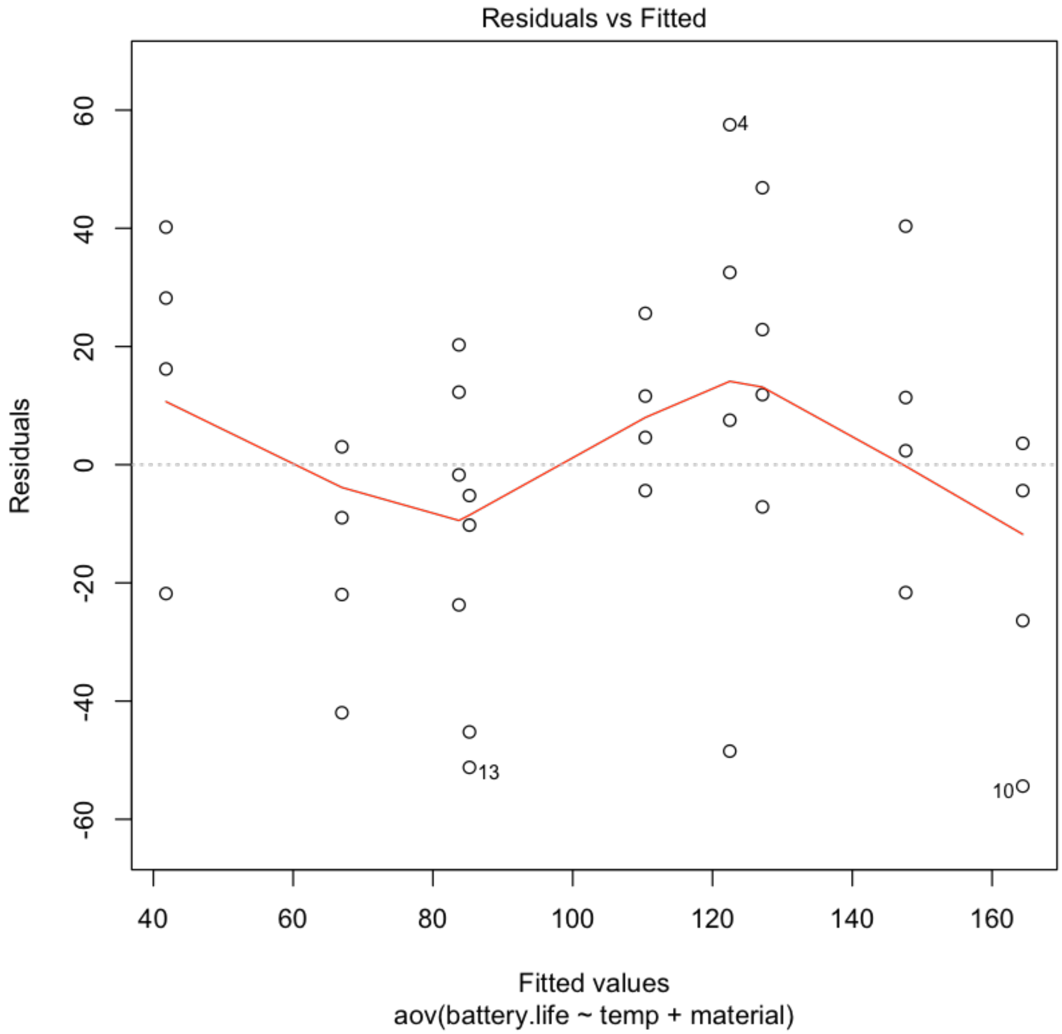
```
In [28]: no.int.residuals <- residuals(battery.no.int.aov)

         no.int.fitted <- fitted(battery.no.int.aov)
```

```
In [29]: plot(no.int.fitted, no.int.residuals)
```

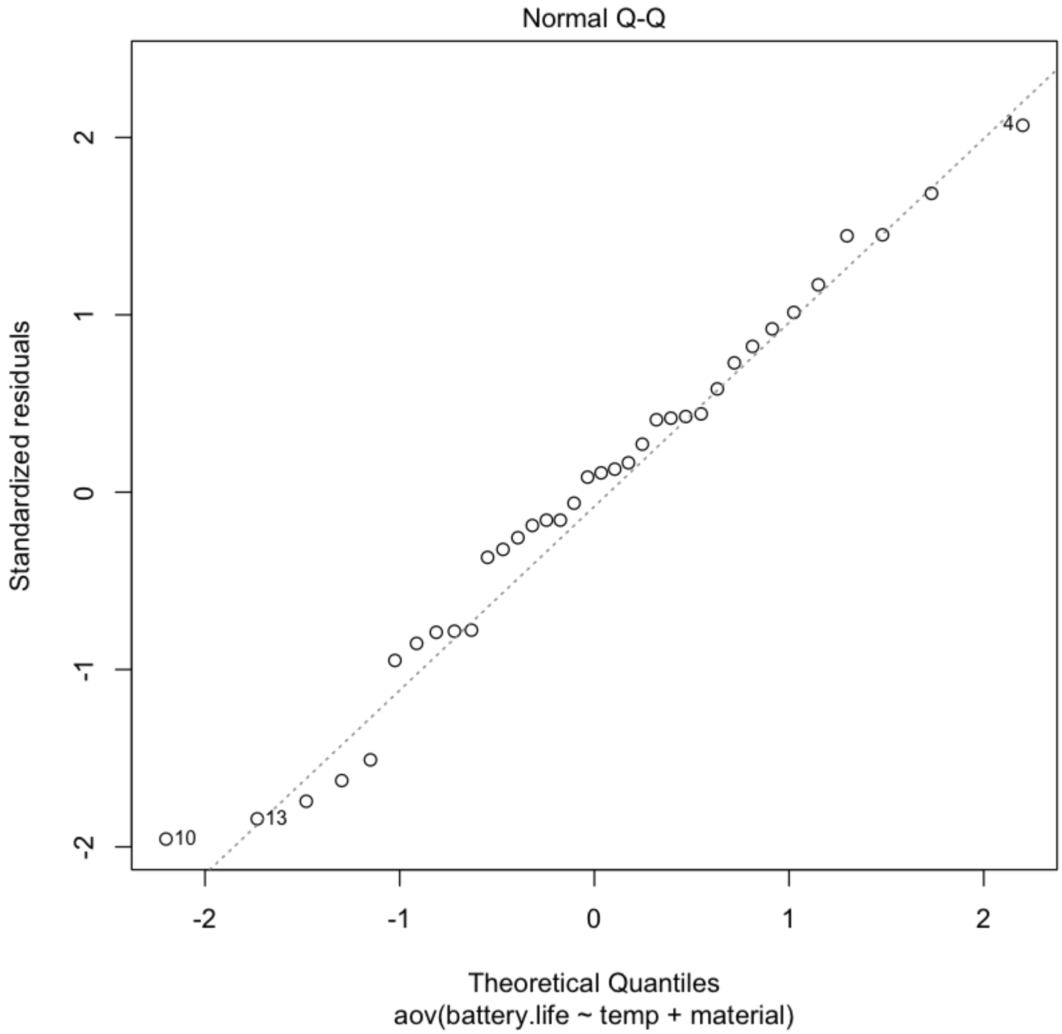


0.12 Note, you can get the exact same residuals Vs fitted values plot from just plotting the AOV object

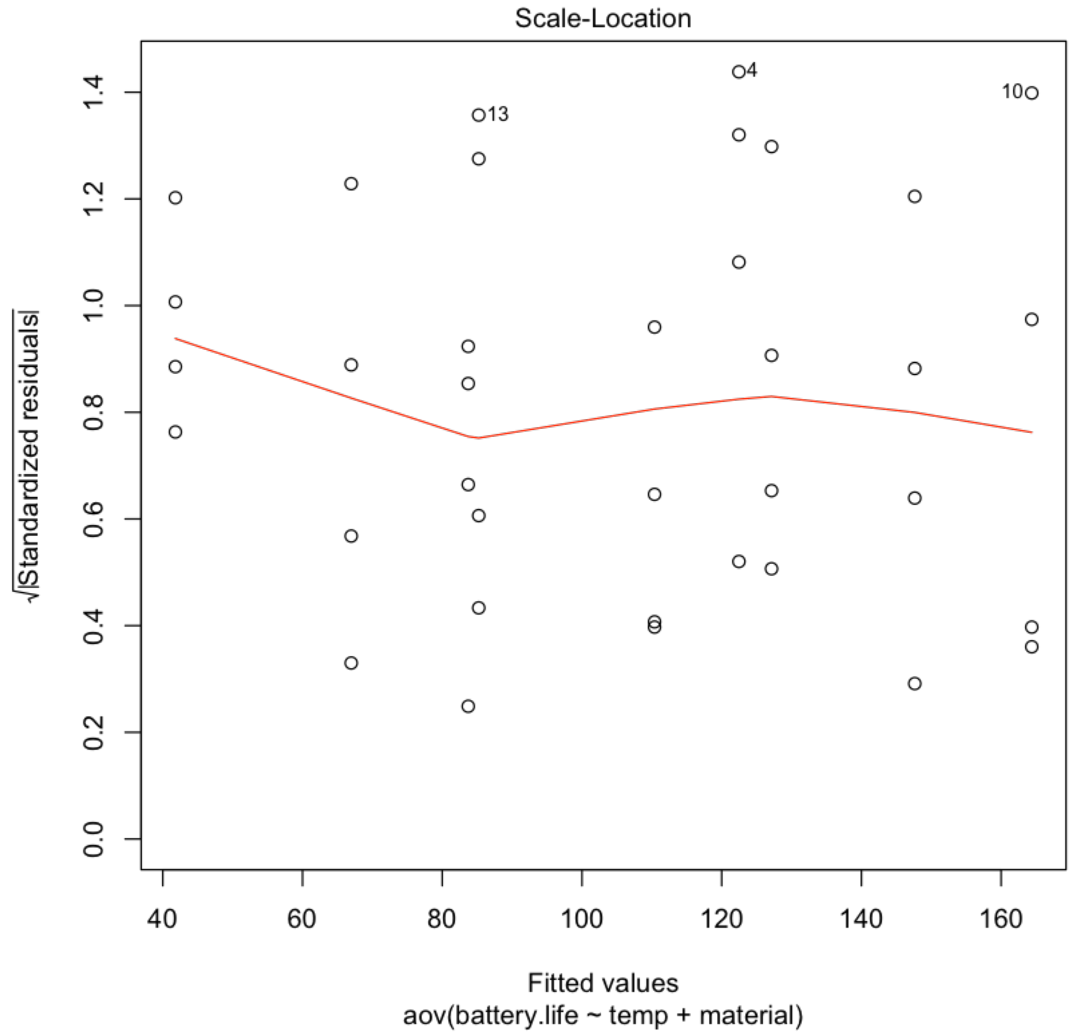
```
In [30]: plot(battery.no.int.aov)
```



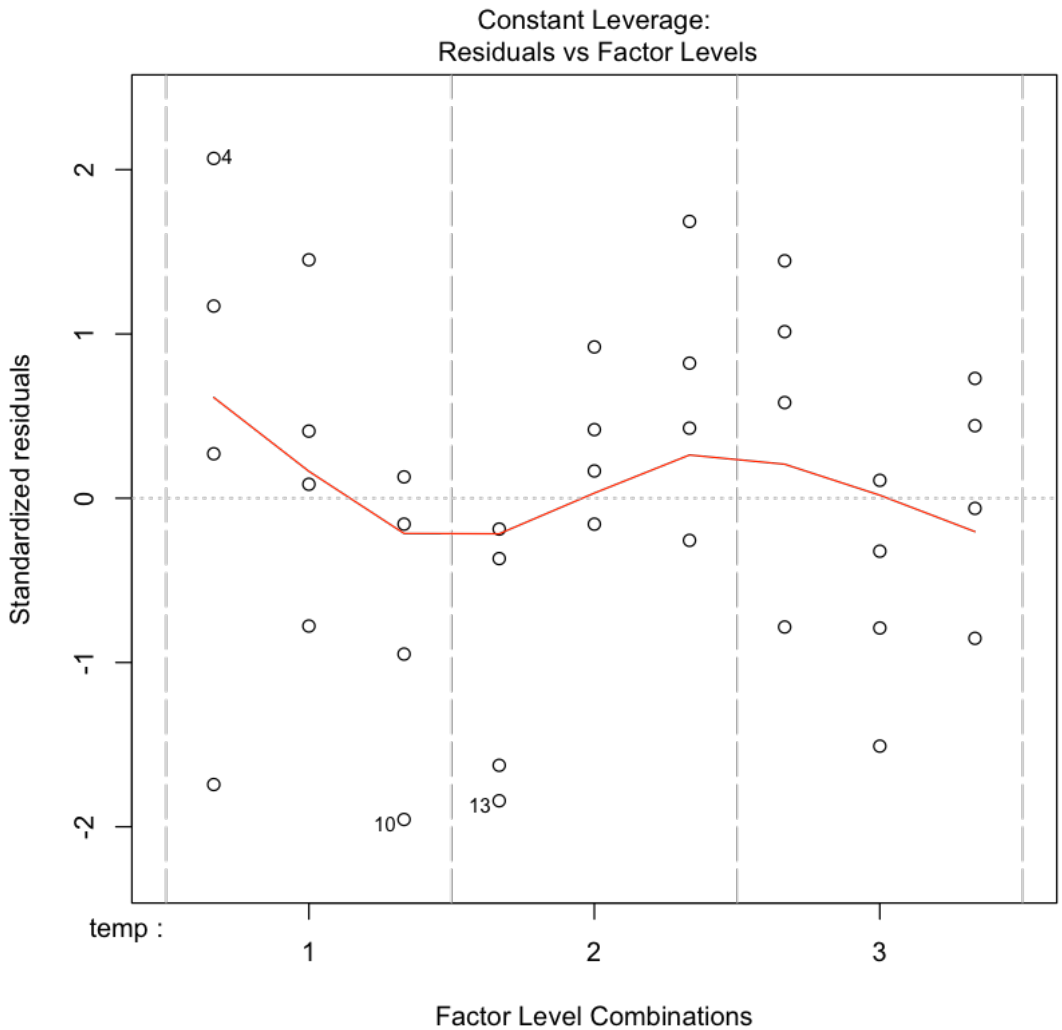
Residuals vs Fitted
Residuals
60
40
20
0
-20
-40
-60
4
13
10
40
60
80
100
120
140
160
Fitted values
aov(battery.life ~ temp + material)



Normal Q-Q
Standardized residuals
Theoretical Quantiles
aov(battery.life ~ temp + material)
-2
-1
0
1
2
10
13
4



Scale-Location
√|Standardized residuals|
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
4
13
10
40
60
80
100
120
140
160
Fitted values
aov(battery.life ~ temp + material)



Based on the model adequecy test in the residuals against fitted values plot, we see a rise and fall and rise and fall in the residuals, and hence there's non-constant error variance that is being assumed as part of the linear model, therefore we reject the model, favouring the model including the model **including interaction terms**.

CG Note: The QQ plot actually shows the non-interacting model to be closer to the assumption of normality than the interacting model...