Data Science 1
Experimental Design and Analysis of Variance
2-Factor Factorial Design

Ann Maharaj

Experimental Design: 2-Factor Factorial Design

Introduction

- In the previous topics we learnt about the:
  - Completely randomised design which is a one-factor factorial design, where we used a one-way ANOVA to test for differences in means several independent populations.
  - Randomised block design which involves one factor and a block, where we used a two-way ANOVA to test for differences in means of several related populations.

- In this topic we will extend the analysis of variance to the two-factor factorial design in which two factors are simultaneously evaluated.

Introduction

- A two-factor factorial design is an experimental design in which data is collected for all possible combinations of the levels of the two factors of interest.
- If equal sample sizes are taken for each of the possible factor combinations then the design is a balanced two-factor factorial design.
- A balanced $a \times b$ factorial design is one for which there are a levels of Factor A, b levels of Factor B, and r independent replications taken at each of the $a \times b$ treatment combinations. The design size is $n = abr$.
- The effect of a factor is defined to be the average change in the response associated with a change in the level of the factor. This is usually called a main effect.
- If the average change in response across the levels of one factor are not the same at all levels of the other factor, then we say there is an interaction between the factors.

Introduction

- The two-factor factorial design is implemented by a two-way ANOVA which operates in the same way as the one-way ANOVA except that additional independent variables are examined. The method is referred to as two-way ANOVA with replications.
- Each independent variable or factor must be evaluated at two or more levels, viz., treatments . In a two-factor design, each case has been randomly assigned to only one of the different factor levels of each of the independent variables.
- Each of the different cells represents the unique combinations of the levels of the two factors which are the treatment combinations.

Introduction

An engineer tests three plate materials for a new battery at three temperature levels ($-9^o C, 21^o C, 68^o C$). Four batteries (replicates) are tested with each combination of plate material and temperature. All 36 tests are run in random order and the lifetimes in hours are recorded. What effect do material type and temperature have on the life of a battery?

Table: Lifetime in hours for Battery Design Experiment

	Temperature Level		
	1	2	3
Material Type			
1	130	34	20
	74	80	82
	155	40	70
	180	75	58
2	150	136	25
	188	122	70
	159	106	58
	126	115	45
3	138	174	96
	168	120	104
	110	150	82
	160	139	60

- Two independent variables: Material Type, Temperature Level
- Factor A: Material Type; a = 3 treatment levels
- Factor B: Temperature Level: b = 3 treatment levels
- Number of replicates: r = 4
- Experimental units: 36 batteries
- Response (dependent variable): Lifetime in hours

Data Structure

Table: Data Structure for 2-Factor Factorial Design

	Factor B						**Factor A**
	Treatment 1	Treatment 2				Treatment b	**Means**
Factor A Treatment 1	y_{111}	y_{121}	.	.	.	y_{1b1}	
	y_{112}	y_{122}	.	.	.	y_{1b2}	
	.	.	.	.	.	.	
	.	.	.	.	.	.	
	y_{11r}	y_{12r}	.	.	.	y_{1br}	
	$\bar{y}_{11.}$	$\bar{y}_{12.}$	.	.	.	$\bar{y}_{1b.}$	$\bar{y}_{1..}$
Treatment 2	y_{211}	y_{221}	.	.	.	y_{2b1}	
	y_{212}	y_{222}	.	.	.	y_{2b2}	
	.	.	.	.	.	.	
	.	.	.	.	.	.	
	y_{21r}	y_{22r}	.	.	.	y_{2br}	
	$\bar{y}_{21.}$	$\bar{y}_{22.}$				$\bar{y}_{2b.}$	$\bar{y}_{2..}$
	.	.	.	.	.	.	
	.	.	.	.	.	.	
	.	.	.	.	.	.	
Treatment a	y_{a11}	y_{a21}	.	.	.	y_{ab1}	
	y_{a12}	y_{a22}	.	.	.	y_{ab2}	
	.	.	.	.	.	.	
	.	.	.	.	.	.	
	y_{a1r}	y_{a2r}	.	.	.	y_{abr}	
	$\bar{y}_{a1.}$	$\bar{y}_{a2.}$				$\bar{y}_{ab.}$	$\bar{y}_{a..}$
Factor B Means	$\bar{y}_{.1.}$	$\bar{y}_{.2.}$				$\bar{y}_{.b.}$	**Grand Mean** $\bar{y}_{...}$

Data Structure

Factor A x Factor B ij^{th} Treatment Combination Mean:

$$\overline{y}_{ij.} = \frac{1}{r}\sum_{k=1}^{r} y_{ijk}$$

i^{th} Factor A (Row) Mean:

$$\overline{y}_{i..} = \frac{1}{b}\sum_{j=1}^{b} \overline{y}_{ij.}$$

j^{th} Factor B (Column) Mean:

$$\overline{y}_{.j.} = \frac{1}{a}\sum_{i=1}^{a} \overline{y}_{ij.}$$

Grand mean

$$\overline{y}_{...} = \frac{1}{a}\sum_{i=1}^{a} \overline{y}_{i..} = \frac{1}{b}\sum_{j=1}^{b} \overline{y}_{.j.}$$

Partition of Total Variation

Given the basic decomposition

$$(y_{ijk} - \overline{y}_{...}) = (\overline{y}_{i..} - \overline{y}_{...}) + (\overline{y}_{.j.} - \overline{y}_{...}) + (\overline{y}_{ij.} - \overline{y}_{i..} - \overline{y}_{.j.} + \overline{y}_{...}) + (y_{ijk} - \overline{y}_{ij.})$$

the following decomposition in sums of square can be obtained

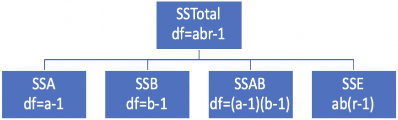
$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}(y_{ijk} - \overline{y}_{...})^2 = br\sum_{i=1}^{a}(\overline{y}_{i..} - \overline{y}_{...})^2 + ar\sum_{i=1}^{b}(\overline{y}_{.j.} - \overline{y}_{...})^2$$
$$+r\sum_{j=1}^{a}\sum_{i=1}^{b}(\overline{y}_{ij.} - \overline{y}_{i..} - \overline{y}_{.j.} + \overline{y}_{...})^2$$
$$+\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}(y_{ijk} - \overline{y}_{ij.})^2$$

$\Rightarrow$ Total SS = Factor A SS + Factor B SS + Interaction(AB) SS + Residual or Error SS

That is, the total variation is partitioned into variation due to Factor A, Factor B, interaction between Factors A and B, and due to the residual or error variation.

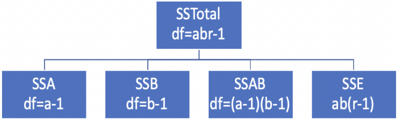
$$SSTotal = SSA + SSB + SSAB + SSE$$

Partition of Total Variation



- Total Sum of Squares: $\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}(y_{ijk}-\overline{y}_{...})^2$
  has degree of freedom $df = abr - 1$
  SSTotal is a measure of the dispersion of all individual data values in relation to the mean $\overline{y}_{...}$ (grand mean) of the entire data set.
- Factor A Sum of Squares: $br\sum_{i=1}^{a}(\overline{y}_{i..}-\overline{y}_{...})^2$
  has degree of freedom $df = a - 1$
  SSA is a measure of dispersion of the Factor A means from the grand mean. It is the variation due to Factor A.
- Factor B Sum of Squares: $ar\sum_{i=1}^{b}(\overline{y}_{.j.}-\overline{y}_{...})^2$
  has degree of freedom $df = b - 1$
  SSB is a measure of dispersion of the Factor B means from the grand mean. It is the variation due to Factor B.

Partition of Total Variation



- Interaction AB Sum of Squares: $SSAB = r\sum_{j=1}^{a}\sum_{i=1}^{b}(\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2$ has degree of freedom $df = (a-1)(b-1)$
  SSAB is a measure of dispersion of interaction means from the grand mean. It is the variation due to the interaction between A and B.

- Error Sum of Squares: $SSE = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}(y_{ijk} - \bar{y}_{ij.})^2$ has degrees of freedom $df = ab(r-1)$.
  SSE is an aggregated measure of how much data values vary from the interaction means of Factors A and B. It is the variation due to random sampling (unexplained variation).

Mean Squares

Mean Square = Sum of Squares/df

MSA = SST/(a-1)

MSB = SSB/(b-1)

MSAB = SSB/(a-1)(b-1)

MSE = SSE/(ab(r-1)

Tests for Factor A treatment effects, Factor B treatment effects, and the interaction of the effects of Factors A and B treatments can now be performed by comparing the corresponding mean square with the yardstick of the residual mean square *MSE* which can be shown to be an unbiased estimator of the common population variance σ^2.

Population Model for a 2-Factor Factorial Design

- To implement a formal statistical test for Factor A and B treatment effects, as well as the interaction of the effects of Factors A and B treatments, we require a population model for the experiment.

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$
$$i = 1, 2, ..., a, \ j = 1, 2, ..., b$$

$\mu = \frac{\sum_{i=1}^{a} \mu_{i..}}{a} = \frac{\sum_{j=1}^{b} \mu_{.j.}}{b}$ = overall mean

$\alpha_i = (\mu_{i..} - \mu)$: i^{th} Factor A effect, $\sum_{i=1}^{a} \alpha_i = 0$

$\beta_j = (\mu_{.j.} - \mu)$: j^{th} Factor B effect, $\sum_{j=1}^{b} \beta_j = 0$

$(\alpha\beta)_{ij} = (\mu_{ij.} - \mu)$: $(i,j)^{th}$ AB interaction effect,
$\sum_{i=1}^{a} (\alpha\beta)_{ij} = 0$ for $i = 1, 2, ..., a$,
$\sum_{j=1}^{b} (\alpha\beta)_{ij} = 0$ for $j = 1, 2, ..., b$

ε_{ijk}: random error of the k^{th} observation from the $(i,j)^{th}$ cell.

- Assume ε_{ijk} are independently and identically distributed as $N(0, \sigma^2)$ and that the effects are fixed.

Hypothesis Testing

- Testing Factor A treatment effects:
  $H_0 : \alpha_1 = \alpha_2 = ... = \alpha_a = 0$
  All a treatment effects are equal to zero which implies there is no Factor A treatment effect and hence the Factor A population means are all equal.

  H_1: Not all the α_i are equal to zero, $i = 1, 2, ..., a$.
  At least one of the Factor A treatment effects is not zero which implies here is a Factor A treatment effect and hence at least one Factor A population mean is different.

  Hence the null and alternative hypotheses can be expressed as
  $H_0 : \mu_{1..} = \mu_{2..} = ... = \mu_{a..}$
  H_1: not all $\mu_{i..}$ are equal; $i = 1, 2, ..., a$.

- To test the null hypothesis, we observe that the Factor A treatment mean square, $MSA = \frac{SSA}{a-1}$
  - is expected to be small when the Factor A population means are all equal.
  - is likely to be large when the means differ markedly.

- It can be shown that under the null hypothesis, $F_A = \frac{MSA}{MSE}$ follows an F-distribution with $(a-1)$ and $ab(r-1)$ degrees of freedom.

Hypothesis Testing

- Testing Factor B treatment effects:
  $H_0 : \beta_1 = \beta_2 = ... = \beta_b = 0$
  All b treatment effects are equal to zero which implies there is no Factor B treatment effect and hence the Factor B population means are all equal.

  H_1: Not all the β_j are equal to zero, $j = 1, 2, ..., b$.
  At least one of the Factor B treatment effects is not zero which implies there is a Factor B treatment effect and hence at least one Factor B population mean is different.

  Hence the null and alternative hypotheses can be expressed as
  $H_0 : \mu_{.1.} = \mu_{.2.} = ... = \mu_{.b.}$
  H_1: not all $\mu_{.j.}$ are equal; $j = 1, 2, ..., b$.

- To test the null hypothesis, we observe that the Factor B treatment mean square, $MSB = \frac{SSB}{b-1}$
  - is expected to be small when the Factor B population means are all equal.
  - is likely to be large when the means differ markedly.

- It can be shown that under the null hypothesis, $F_B = \frac{MSB}{MSE}$ follows an F-distribution with $(b-1)$ and $ab(r-1)$ degrees of freedom.

Hypothesis Testing

- Testing AB interaction effects:
  H_0 All $(\alpha\beta)_{ij} = 0,\ i = 1, 2, ..., a,\ j = 1, 2, ..., b$
  All interaction effects are equal to zero which implies there is no interaction effect and hence the AB interaction population means are all equal.

  H_1: Not all $(\alpha\beta)_{ij} = 0,\ i = 1, 2, ..., a,\ j = 1, 2, ..., b$
  At least one of the AB interaction effects is not zero which implies here is a AB interaction effect and hence at least AB interaction population mean is different.

  Hence the null and alternative hypotheses can be expressed as
  H_0: All $\mu_{ij.}$ are equal, $i = 1, 2, ..., a,\ j = 1, 2, ..., b$
  H_1: Not all $\mu_{ij.}$ are equal.

- To test the null hypothesis, we observe that the AB interaction mean square, $MSAB = \frac{SSAB}{(a-1)(b-1)}$
  - is expected to be small when the AB interaction population means are all equal.
  - is likely to be large when the means differ markedly.

- It can be shown that under the null hypothesis, $F_{AB} = \frac{MSAB}{MSE}$ follows an F-distribution with $(a-1)(b-1)$ and $ab(r-1)$ degrees of freedom.

2-Way ANOVA Table

Table: 2-Way ANOVA Table

Source	Sum of Squares	df	Mean Square	F-Ratio
Factor A	SSA	a-1	$MSA = \frac{SSA}{a-1}$	$F_A = \frac{MST}{MSE}$
Factor B	SSB	b-1	$MSB = \frac{SSB}{b-1}$	$F_B = \frac{MSB}{MSE}$
AB Interaction	$SSAB$	$(a-1)(b-1)$	$MSAB = \frac{SSAB}{(a-1)(b-1)}$	$F_{AB} = \frac{MSAB}{MSE}$
Error	SSE	$ab(r-1)$	$MSE = \frac{SSE}{ab(r-1)}$	
Total	$SSTotal$	$abr-1$		

Because there are two factors each with replications, we refer to this setup as a Two-Way ANOVA with replications.

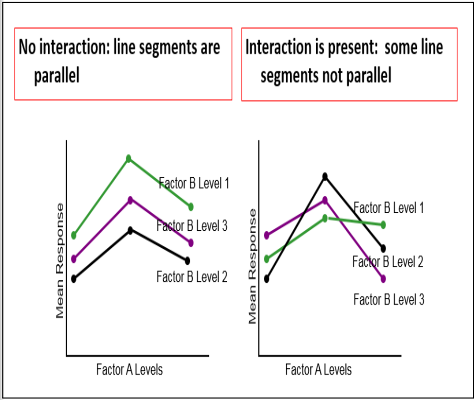
Features of Two-Way ANOVA

- Total = Factor A + Factor B + AB Interaction + Error
- The sums of squares always add up
  - SSTotal = SSA + SSB + SSAB + SSE
- Degrees of freedom always add up.
  - $abr - 1 = (a-1) + (b-1) + (a-1)(b-1) + ab(r-1)$
- The denominators of the F- Ratios are always the same but the numerators are different.

AB Interaction Effects

Testing AB interaction effects:

1. Hypotheses:
   H_0: All $\mu_{ij.}$ are equal, $i = 1, 2, ..., a$, $j = 1, 2, ..., b$
   H_1: Not all $\mu_{ij.}$ are equal.
2. Level of Significance: α
3. Test Statistic: $F_{AB} = \frac{MSAB}{MSE}$
4. P-value: p-value $= P(F > F_{AB})$
5. Decision Rule: Reject H_0 if $p - value < \alpha$
6. Conclusion



AB Interaction Effects

- Parallel lines usually indicate no significant interaction.
- Severe lack of parallelism usually indicates a significant interaction.
- Moderate lack of parallelism suggests a possible significant interaction may exist.
- Statistical significance of an interaction effect depends on the magnitude of the MSE. For small values of the MSE, even small interaction effects (less non-parallelism) may be significant.
- When an AB interaction is large, the corresponding main effects A and B may have little practical meaning.
    - Knowledge of the AB interaction is often more useful than knowledge of the main effects. i.e., the experimenter must examine the levels of one factor, say A, at fixed levels of the other factor to draw conclusions about the main effect of A.
- It is possible to have a significant interaction between two factors while the main effects for both factors are not significant.
    - This would happen when the interaction plot shows interactions in different directions that balance out over one or both factors (such as an X pattern). This type of interaction, however, is uncommon.

Hypothesis Testing

Testing Factor A Treatment Effects

1. Hypotheses:
   $H_0 : \mu_{1..} = \mu_{2..} = ... = \mu_{a..}$
   H_1: not all $\mu_{i..}$ are equal;
   $i = 1, 2, ..., a$.
2. Level of Significance: α
3. Test Statistic: $F_A = \frac{MSA}{MSE}$
4. P-value: p-value $= P(F > F_A)$
5. Decision Rule: Reject H_0 if
   $p - value < \alpha$
6. Conclusion

Testing Factor B Treatment Effects

1. Hypotheses:
   $H_0 : \mu_{.1.} = \mu_{.2.} = ... = \mu_{.b.}$
   H_1: not all $\mu_{.j.}$ are equal;
   $j = 1, 2, ..., b$.
2. Level of Significance: α
3. Test Statistic: $F_B = \frac{MSB}{MSE}$
4. P-value: p-value $= P(F > F_B)$
5. Decision Rule: Reject H_0 if
   $p - value < \alpha$
6. Conclusion

Factor A Multiple Comparisons

- It can be shown that $(\overline{y}_{i..} - \overline{y}_{i'..})$ is normally distributed with a mean of $(\mu_{i..} - \mu_{i'..})$
  and variance $\sigma^2 \left(\frac{1}{br} + \frac{1}{br}\right) = \sigma^2 \left(\frac{2}{br}\right)$, where σ^2 is unknown and is estimated by MSE.

- Hence, it can be shown that *Tukey's Honestly Significant Difference* test statistic is

$$T_{HSD} = \frac{(\overline{y}_{i..} - \overline{y}_{i'..}) - (\mu_{i..} - \mu_{i'..})}{\sqrt{MSE\left(\frac{2}{br}\right)}}$$

  follows a t-distribution with $df = ab(r-1)$.

Factor A Multiple Comparisons

- It can be shown that a $100(1-\alpha)\%$ confidence interval for each of the m_1 number of multiple pairwise differences $\mu_{i..} - \mu_{i'..}$ is given by

$$(\overline{y}_{i..} - \overline{y}_{i'..}) \pm t_{\alpha/2m_1, ab(r-1)} \sqrt{MSE\left(\frac{2}{br}\right)}$$

where $t_{\alpha/2m_1, ab(r-1)}$ is the upper $\alpha/2m_1$ quantile of the t-distribution with $df = ab(r-1)$

- Using this procedure, the probability of all m_1 statements being correct is at least $(1-\alpha)$. Dividing α by $2m_1$ adjusts for the probability of the Type 1 error since m_1 pairwise comparisons are required.

- The p-value of each test can also be obtained by also adjusting for the probability of Type 1 error.

Factor B Multiple Comparisons

- It can be shown that $(\overline{y}_{.j.} - \overline{y}_{.j'.})$ is normally distributed with a mean of $(\mu_{.j.} - \mu_{.j'.})$

  and variance $\sigma^2 \left(\frac{1}{ar} + \frac{1}{ar}\right) = \sigma^2 \left(\frac{2}{ar}\right)$, where σ^2 is unknown and is estimated by MSE.

- Hence, it can be shown that *Tukey's Honestly Significant Difference* test statistic is

$$T_{HSD} = \frac{(\overline{y}_{.j.} - \overline{y}_{.j'.}) - (\mu_{.j.} - \mu_{.j'.})}{\sqrt{MSE \left(\frac{2}{ar}\right)}}$$

  follows a t-distribution with $df = ab(r-1)$.

Factor B Multiple Comparisons

- It can be shown that a $100(1-\alpha)\%$ confidence interval for each of the m_2 number of multiple pairwise differences $\mu_{.j.} - \mu_{.j'.}$ is given by

$$(\overline{y}_{.j.} - \overline{y}_{.j'.}) \pm t_{\alpha/2m_2, ab(r-1)} \sqrt{MSE\left(\frac{2}{ar}\right)}$$

  where $t_{\alpha/2m_2, ab(r-1)}$ is the upper $\alpha/2m_2$ quantile of the t-distribution with $df = ab(r-1)$.

- Using this procedure, the probability of all m_2 statements being correct is at least $(1-\alpha)$. Dividing α by $2m_2$ adjusts for the probability of the Type 1 error since m_2 pairwise comparisons are required.

- The p-value of each test can also be obtained by also adjusting for the probability of Type 1 error.

Assumptions and Diagnostic Checks

Assumptions

1. Random Sampling: the observations should be from random samples drawn from the independent populations of interest.
2. Population Normality: The populations from which the samples are drawn should be normal.
3. Homogeneity of Variance: variances of populations should be equal.

Diagnostic Checks

- We perform the normality checks on the residuals of the estimated model examining a normal Q-Q plot and the Shapiro-Wilk test for normality.
- We perform homogeneity of variances checks on the residuals by examining a residual plot and also using Levene's test.

Levene's Test for Homogeneity of Variance

H_0: All $\sigma^2_{ij.} = \sigma^2$, $i = 1, 2, ..., a, j = 1, 2, ..., b$
H_1: Not all of the $\sigma^2_{ij.}$ are equal,

To test these hypotheses of equal variances:

- Subtract from each $(ijk)^{th}$ combination of Factor A and B treatments, the treatment mean of the corresponding r replications and obtain the absolute value.

$$\left|y_{ijk} - \overline{y}_{ij.}\right| \text{ , } i = 1, 2, ..., a, \ j = 1, 2, ..., b, \ k = 1, 2, ..., r$$

- Perform a one-way ANOVA on these absolute measures.

Example

An engineer tests three plate materials for a new battery at three temperature levels ($-9^o C, 21^o C, 68^o C$). Four batteries (replicates) are tested with each combination of plate material and temperature. All 36 tests are run in random order and the lifetimes in hours are recorded. What effect do material type and temperature have on the life of a battery?

Table: Lifetime in hours for Battery Design Experiment

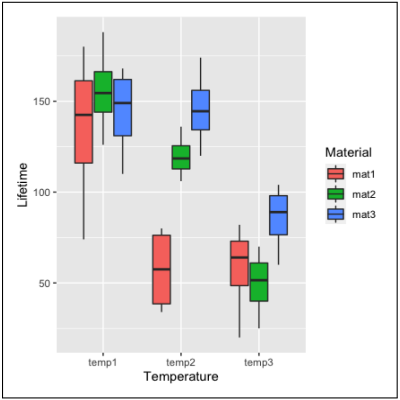
	Temperature Level		
	1	2	3
Material Type			
1	130	34	20
	74	80	82
	155	40	70
	180	75	58
2	150	136	25
	188	122	70
	159	106	58
	126	115	45
3	138	174	96
	168	120	104
	110	150	82
	160	139	60

- Two independent variables: Material Type, Temperature Level
- Factor A: Material Type; a = 3 treatment levels
- Factor B: Temperature Level: b = 3 treatment levels
- Number of replicates: r = 4
- Experimental units: 36 batteries
- Response (dependent variable): Lifetime in hours
- This data is in the file lifetime.csv.

1. What effect do material type and temperature have on the life of a battery? Perform tests at the 5% significance level.
2. If necessary, perform multiple comparison tests.
3. Perform relevant diagnostic checks for normality and homogeneity of variances.

Example

temp1 $= -9^o C$, temp2 $= 21^o C$, temp3 $= 68^o C$



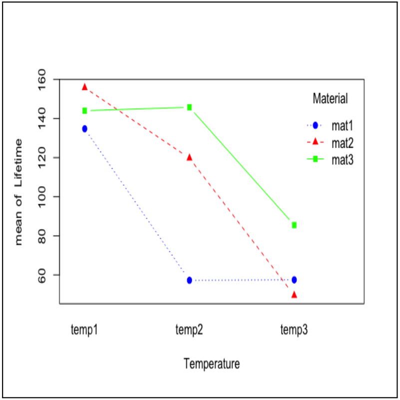
Batteries made from

- Material 1 have greater lifetimes at a very low temperature, compared to medium and very high temperatures for which they are somewhat similar.
- Material 2 have very different lifetimes at the different temperatures with the greatest being at a very low temperature, followed by medium and high temperatures.
- Material 3 have greater lifetimes which are somewhat similar at very low and medium temperatures, but much less so at a high temperature.
- Hence it does appear that material type and temperature have on an effect the life of a battery but this will have to be tested.

Example - Part 1

Factor A: Material Type
Factor B: Temperature Level

- We first assess if there is any interaction between the factors and then test for it.
- The non-parallelism of the segments together and some crossing over is a clear indication of interaction.



Example - Part 1

AB Interaction Effects (Between Material Type and Temperature Level):

1. Hypotheses:
   H_0: All $\mu_{ij.}$ are equal, $i = 1, 2, 3$, $j = 1, 2, 3$
   H_1: Not all $\mu_{ij.}$ are equal.
2. Level of Significance: $\alpha = 0.05$
3. Test Statistic: $F_{A*B} = 3.560$
4. P-value: p-value $= 0.019$
5. Decision Rule: Reject H_0 if $p - value < 0.05$
6. Since $p - value < 0.05$, reject H_0 at the 5% level of significance.

```
> #Two-way ANOVA
> life.aov<- aov(Lifetime~Material*Temperature)
> summary(life.aov)
                     Df Sum Sq Mean Sq F value   Pr(>F)
Material              2  10684    5342   7.911  0.00198 **
Temperature           2  39119   19559  28.968 1.91e-07 ***
Material:Temperature  4   9614    2403   3.560  0.01861 *
Residuals            27  18231     675
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Conclude there is a significant effect between the material the battery is made of and the temperature at which it is tested.

Example - Part 1

```
> #Two-way ANOVA
> life.aov<- aov(Lifetime~Material*Temperature)
> summary(life.aov)
                     Df Sum Sq Mean Sq F value   Pr(>F)
Material              2  10684    5342   7.911 0.00198 **
Temperature           2  39119   19559  28.968 1.91e-07 ***
Material:Temperature  4   9614    2403   3.560 0.01861 *
Residuals            27  18231     675
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Factor A Treatment Effects (Difference between Material Types):

1. Hypotheses:
   $H_0 : \mu_{1.} = \mu_{2.} = \mu_{3.}$
   H_1: both all $\mu_{i.}$ are equal; $i = 1, 2, 3,$
2. Level of Significance: $\alpha = 0.05$
3. Test Statistic: $F_A = 7.911$
4. P-value: $p - value = 0.002$
5. Decision Rule: Reject H_0 if $p - value < 0.05$
6. SInce the $p - value < 0.05$, reject H_0 at the 5% level and down to the 0.2% level of significance. Conclude that here are significant differences in the mean lifetimes of the batteries made from the three material types.

Factor B Treatment Effects (Difference between Temperature Levels):

1. Hypotheses:
   $H_0 : \mu_{.1} = \mu_{.2} = \mu_{.3}$
   H_1: both all $\mu_{.j}$ are equal; $j = 1, 2, 3.$
2. Level of Significance: $\alpha = 0.05$
3. Test Statistic: $F_B = 28.968.00$
4. P-value: $p - value = \approx 0.000$
5. Decision Rule: Reject H_0 if $p - value < 0.05$
6. Since the $p - value < 0.05$, reject H_0 at the 5% level and indeed any reasonable level of significance. Conclude that here are significant differences in the mean lifetimes of the batteries tested at the three temperature levels.

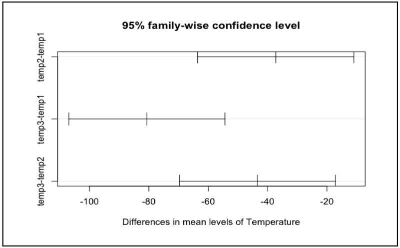
Example - Part 2

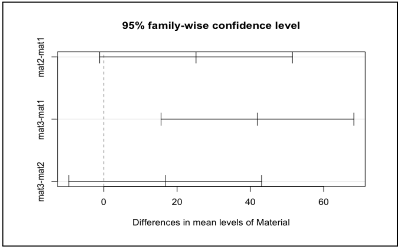
```
> life.tukey <- TukeyHSD(life.aov)
> life.tukey
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = Lifetime ~ Material * Temperature)

$Material
              diff       lwr      upr     p adj
mat2-mat1 25.16667 -1.135677 51.46901 0.0627571
mat3-mat1 41.91667 15.614323 68.21901 0.0014162
mat3-mat2 16.75000 -9.552344 43.05234 0.2717815

$Temperature
                 diff        lwr       upr     p adj
temp2-temp1 -37.25000  -63.55234 -10.94766 0.0043788
temp3-temp1 -80.66667 -106.96901 -54.36432 0.0000001
temp3-temp2 -43.41667  -69.71901 -17.11432 0.0009787
```

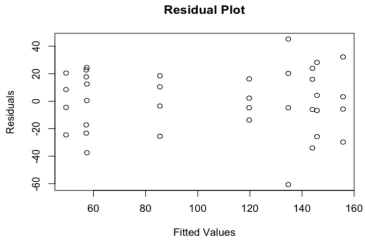




- There are significant difference in the mean lifetimes of batteries for all pairs of temperature levels.
- There is a significant difference in mean lifetimes of batteries made from material types 1 and 3.

Example - Part 3

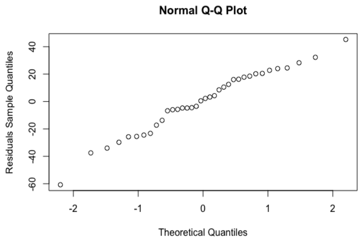
```
> #residuals and fitted values of the model
> res  <- life.aov$residuals
> fit <- life.aov$fitted.values
> #residual plot to assess constant variance assumption
> plot(fit,res, main="Residual Plot",xlab= "Fitted Values", ylab="Residuals")
```




```
> #Test constant variance assumption
> leveneTest(Lifetime~Material*Temperature)
Levene's Test for Homogeneity of Variance (center = median)
      Df F value Pr(>F)
group  8  0.7996 0.6081
      27
```

- The scatterplot of residuals versus fitted values displays no particular pattern, an indication of constant variances.
- The results of Levene's test indicates that the assumption of constant variances is valid.

Example - Part 3




```
> #assess normality assumption
> qqnorm(res, ylab = "Residuals Sample Quantiles")
> shapiro.test(res)

        Shapiro-Wilk normality test

data:  res
W = 0.97606, p-value = 0.6117
```

- The Normal QQ plot does not show much deviation from the diagonal.
- The Shapiro-Wilk test indicates that the assumption of normality of the errors is valid.