Final Report

Eghorieta

12/1/2021

Project For Design of Experments

For our Design of Experiment Project, we created three different parts which looked at three different design of experiments designs. The three different experiment designs that we preformed were a completely randomized design , a factorial design , and 2^4 factorial design.

Completely Randomized Design

```
##
##      Balanced one-way analysis of variance power calculation
##
##               k = 3
##               n = 52.55574
##               f = 0.2357023
##       sig.level = 0.05
##           power = 0.75
##
## NOTE: n is number in each group
```

For this experiment we required 53 samples for each of the 3 different treatment levels. Which resulted in taking 159 total samples.

Layout of Complete Randomized designs

In this experiment, the 3 different treatments are represented by colors yellow, green and blue. the color blue represents the red ball, the color yellow represents the yellow ball and color green represents the green ball that we used in the actual experiment.

Completely Randomized Design

Plot	Replication	Color of Ball
101	1	blue
102	2	blue
103	3	blue
104	1	yellow
105	1	green
106	2	green
107	4	blue

Plot	Replication	Color of Ball
158	22	green
159	21	blue
249	49	yellow
250	50	yellow
251	52	green
252	53	green
253	51	yellow
254	52	yellow
255	50	blue
256	51	blue
257	52	blue
258	53	blue
259	53	yellow

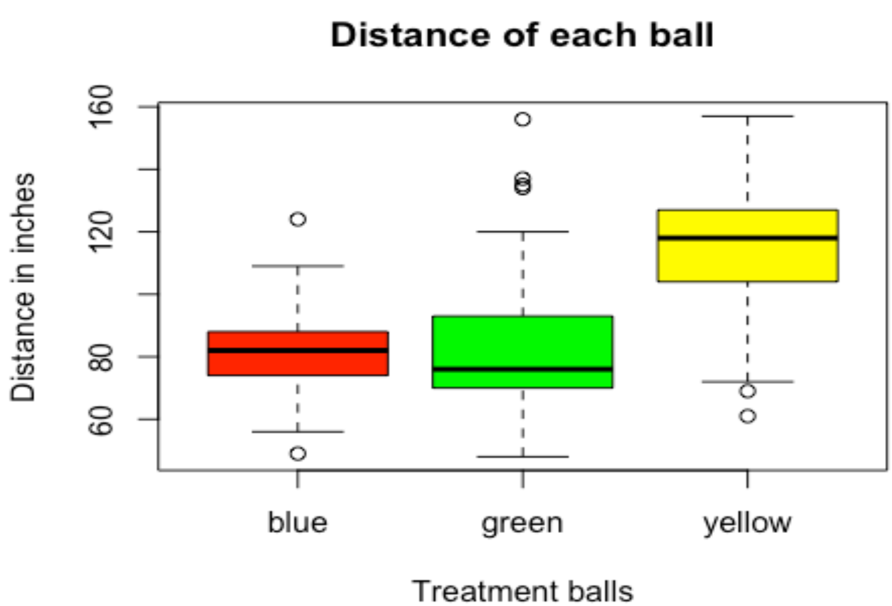
Above is a layout of how we collected the samples for each treatment observation. We saved it in a csv file and used github to read the data into R for further analysis.

Hypothesis test

H$_o$: $\mu_1 = \mu_2 = \mu_3$ - Null Hypothesis

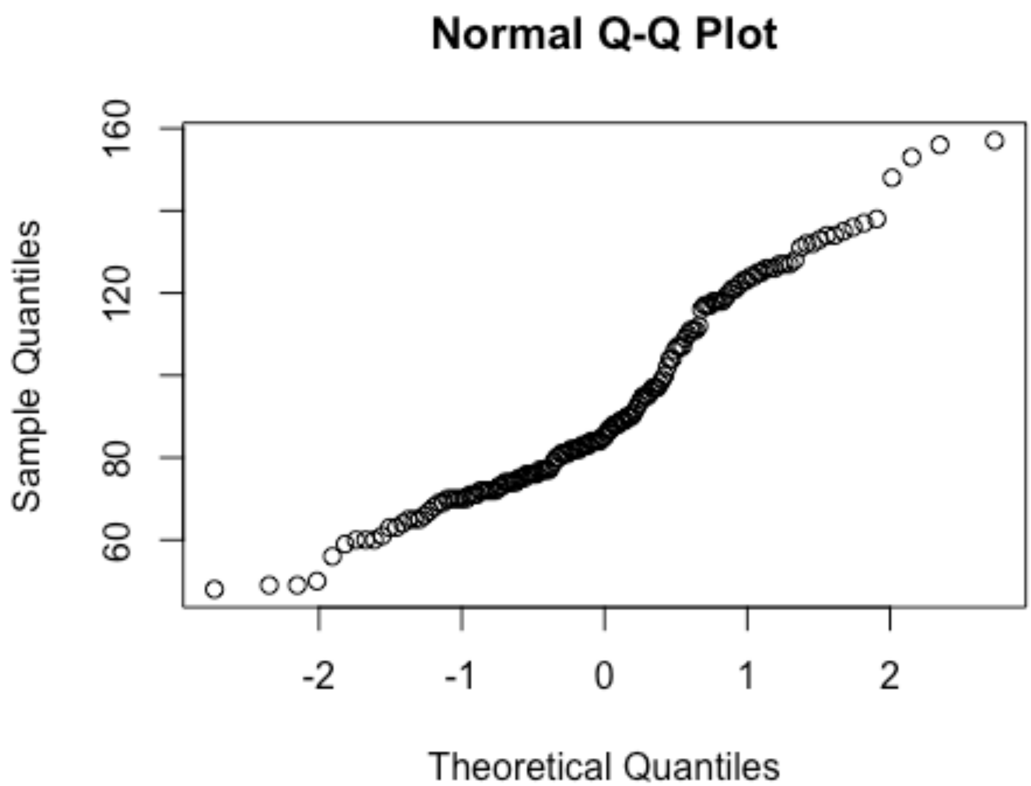
H$_a$: *At least 1 differs* - Alternative Hypothesis

Boxplot of the experiment



The boxplot reveals that the variation between the red ball, green ball and yellow ball are equal.

Testing normality

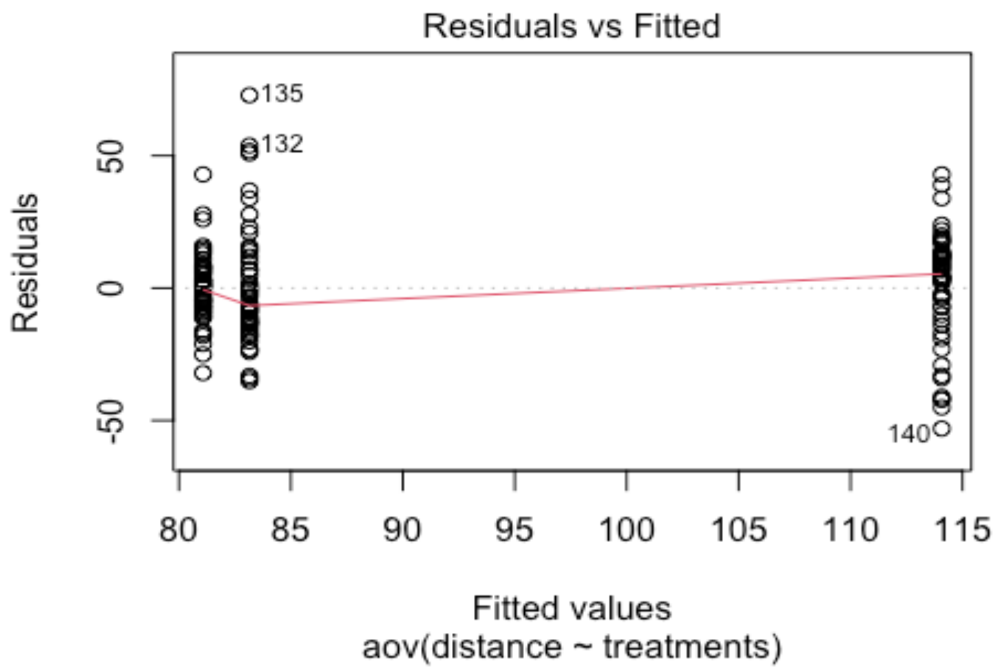


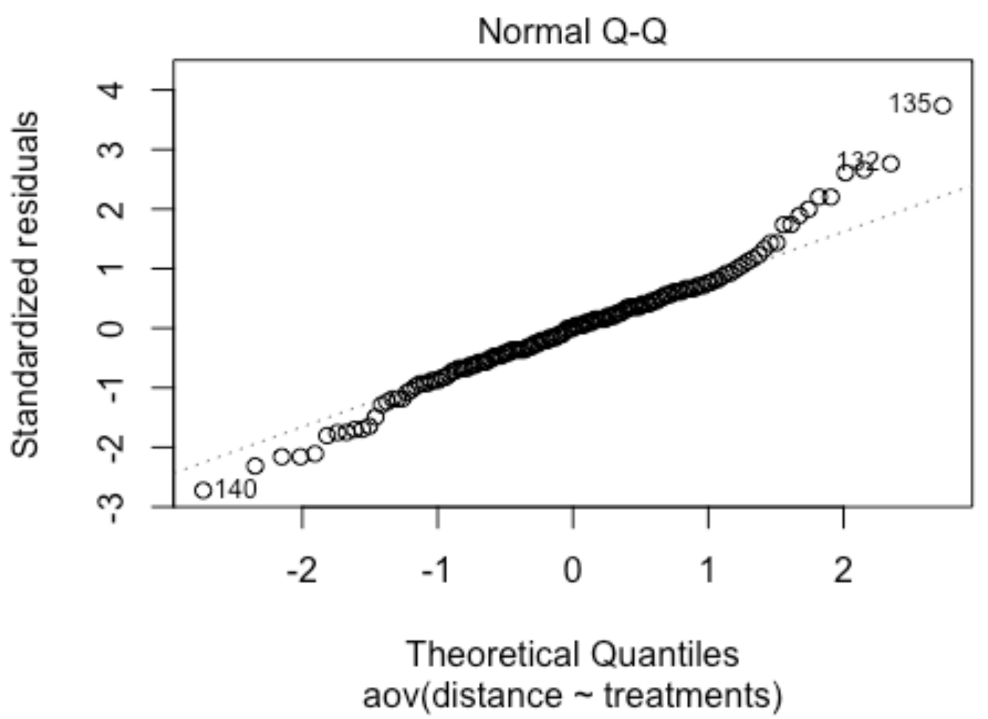
The data looks normally distributed with little presence of outliers at the high extreme values of the distance The outliers might be due to excessive force that was applied to the launching process, the ball landing twice , and a misreading of landing position.

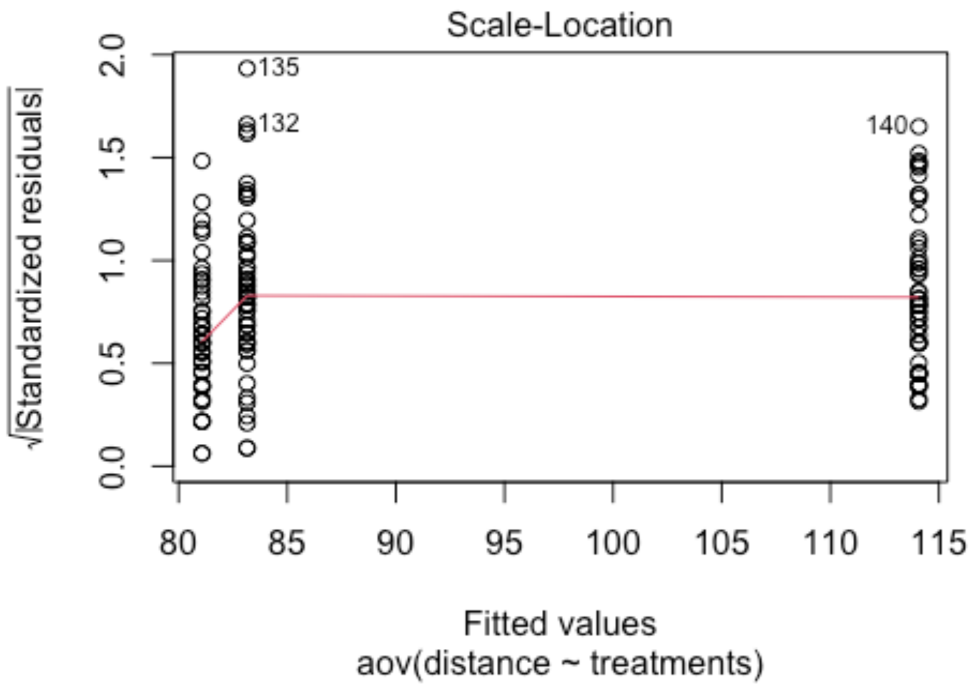
Analysis of variance

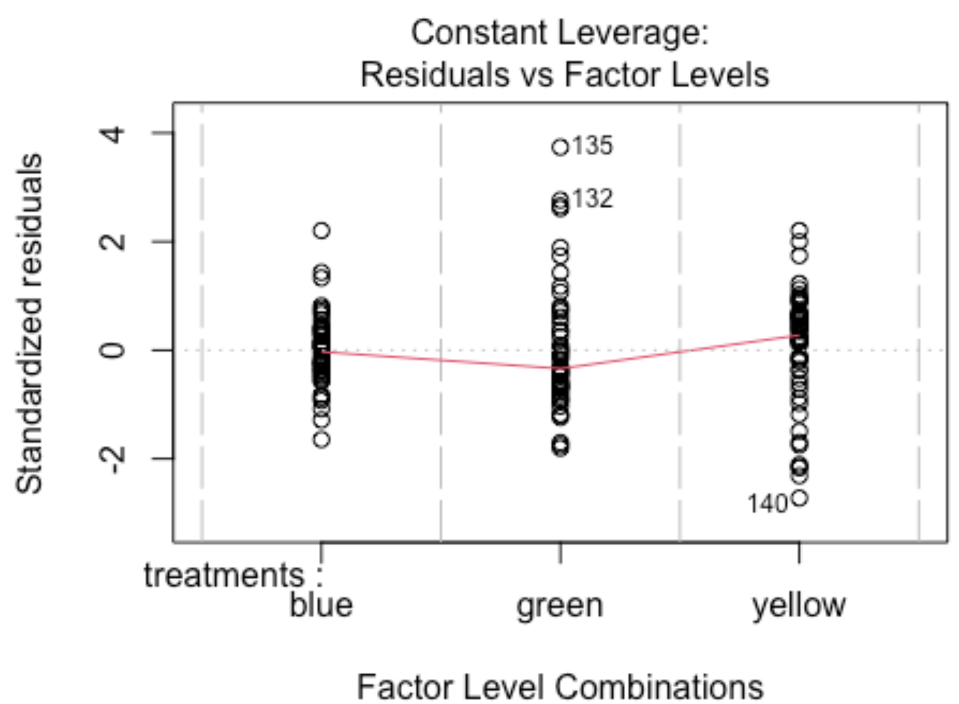
```
##              Df Sum Sq Mean Sq F value Pr(>F)
## treatments    2  36210   18105   46.77 <2e-16 ***
## Residuals   156  60392     387
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

From the result $\mathbf{f_o}$ is **0.783** with a corresponding **p-value** of **0.465** is significantly greater than α = **0.05**. Therefore we fail to reject $\mathbf{H_o}$ that the means are equal, and conclude that none of the means are different.


Residuals vs Fitted
Residuals
Fitted values
aov(distance ~ treatments)
135
132
140


Normal Q-Q
Standardized residuals
Theoretical Quantiles
aov(distance ~ treatments)
135
132
140


Scale-Location
√|Standardized residuals|
135
132
140
2.0
1.5
1.0
0.5
0.0
80
85
90
95
100
105
110
115
Fitted values
aov(distance ~ treatments)


Constant Leverage:
Residuals vs Factor Levels
Standardized residuals
135
132
140
4
2
0
-2
treatments :
blue
green
yellow
Factor Level Combinations

Conclusion

There seems to be nothing unusual about the plots except for the few outliers as the spread of the data looks constant across all treatment balls

Facotorial Design

Null and Alternative Hypotheses

H₀: $\alpha_i = 0$ - Null Hypothesis

Hₐ: $\alpha_i \neq 0$ - Alternative Hypothesis

H₀: $\beta_i = 0$ - Null Hypothesis

Hₐ: $\beta_i \neq 0$ - Alternative Hypothesis

H₀: $\alpha\beta_{ij} = 0$ - Null Hypothesis

Hₐ: $\alpha\beta_{ij} \neq 0$ - Alternative Hypothesis

Level of Significance

α = **0.05**

Model Equation

$$y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \epsilon_{ijk}$$

Proposed Layout with a Randomized Run Order

```
##    plots r A B
## 1    101 1 1 2
## 2    102 2 1 2
## 3    103 1 2 3
## 4    104 1 2 1
## 5    105 1 1 1
## 6    106 2 2 1
## 7    107 1 1 3
## 8    108 1 2 2
## 9    109 3 2 1
## 10   110 2 1 3
## 11   111 3 1 2
## 12   112 2 2 3
## 13   113 2 2 2
## 14   114 3 2 2
## 15   115 3 2 3
## 16   116 2 1 1
## 17   117 3 1 1
## 18   118 3 1 3
```

In the layout, factor A(Pin.Location) represents Pin Elevation and it has levels 1 and 2 for settings 1 and 3 respectively. factor B(Angle) represents the Release Angle with levels 1,2 and 3 for corresponding angles 110, 140 and 170 degrees. Number of replications is 3 which gives a total of 18 observations in the experiment

Collected Data on Proposed Layout

```
##    Replication Pin.Location Angle Distance...Inches.
## 1            1            1   140                 25
## 2            2            1   140                 35
## 3            1            3   170                 55
## 4            1            3   110                 32
## 5            1            1   110                 24
## 6            2            3   110                 23
## 7            1            1   170                 48
## 8            1            3   140                 36
## 9            3            3   110                 24
## 10           2            1   170                 56
## 11           3            1   140                 37
## 12           2            3   170                 61
## 13           2            3   140                 52
## 14           3            3   140                 48
## 15           3            3   170                 72
## 16           2            1   110                 30
## 17           3            1   110                 26
## 18           3            1   170                 33
```

Testing the Hypotheses

```
## Analysis of Variance Table
##
## Response: BungeeEx$Distance...Inches.
##                                       Df  Sum Sq Mean Sq F value    Pr(>F)
## BungeeEx$Pin.Location                  1  440.06  440.06  3.5616 0.1997555
## BungeeEx$Angle                         2 2305.33 1152.67 19.4817 0.0001704
***
## BungeeEx$Pin.Location:BungeeEx$Angle  2  247.11  123.56  2.0883 0.1666387
## Residual                              12  710.00   59.17
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Firstly, we tested the interaction hypothesis that the pin location and the angle had an effect on the shooting distance. If we failed to reject the interaction null hypothesis, we tested the main effects the pin location and angle effects on the distance.

From the interaction result, interaction effects has $\mathbf{f_0}$ value is **2.0883** with a corresponding **p-value** of **0.1666387 >0.05**. Since **0.1666387 >0.05**, we failed to reject the interaction null hypothesis that the interaction between pin location and the angle have an effect on the shooting distance.

The next section we removed the interaction effect and tested the main effects.

Model Equation

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

```
model<-aov(BungeeEx$Distance...Inches.~BungeeEx$Pin.Location+BungeeEx$Angle)
gad(model)
```

```
## Analysis of Variance Table
##
## Response: BungeeEx$Distance...Inches.
##                       Df  Sum Sq Mean Sq F value   Pr(>F)
## BungeeEx$Pin.Location  1  440.06  440.06  6.4368 0.023703 *
## BungeeEx$Angle         2 2305.33 1152.67 16.8605 0.000187 ***
## Residual              14  957.11   68.37
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

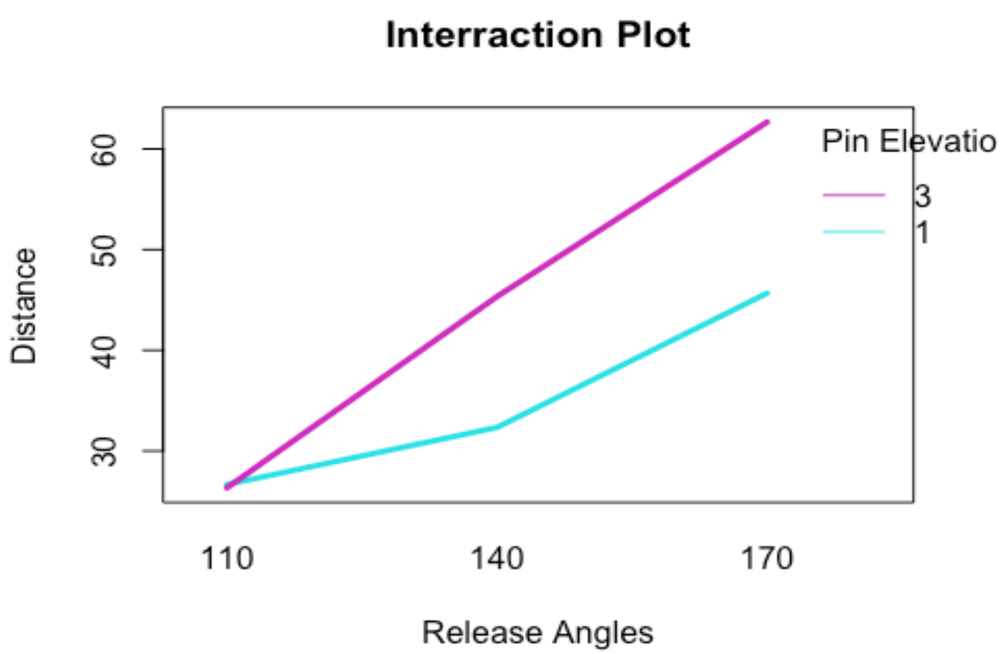
From the pin location result $\mathbf{f_o}$ value is **6.4368** which corresponds to a **p-value** of **0.023703**. The angle result $\mathbf{f_o}$ value is **16.8605** which corresponds to a **p-value** of **0.000187**

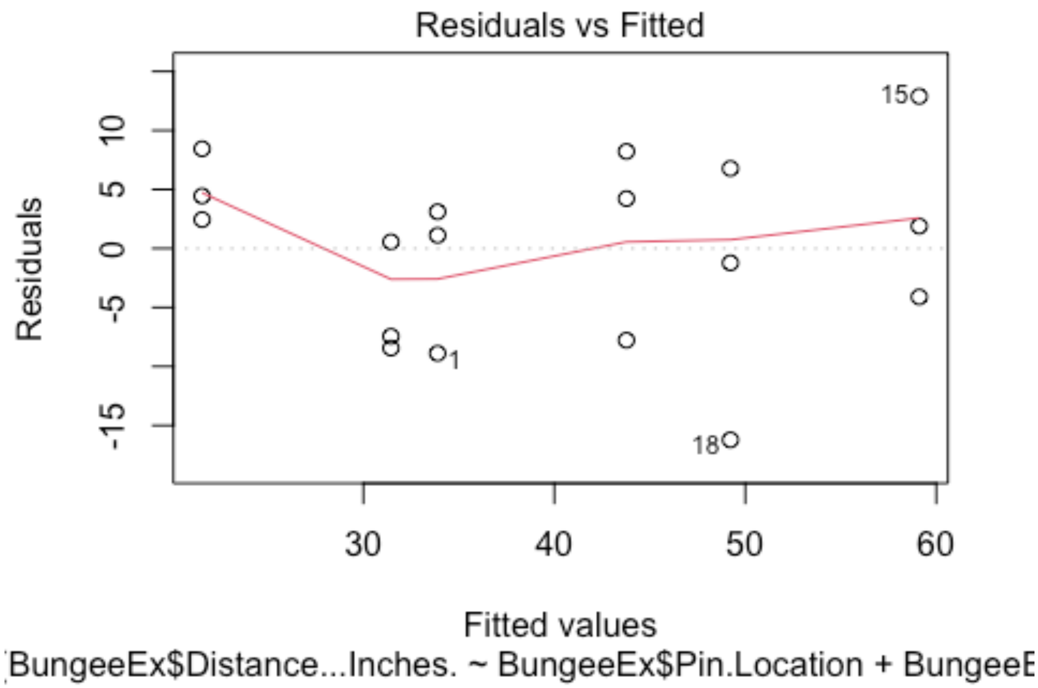
We concluded that the pin location and angle have an effect on the shooting distance of the ball

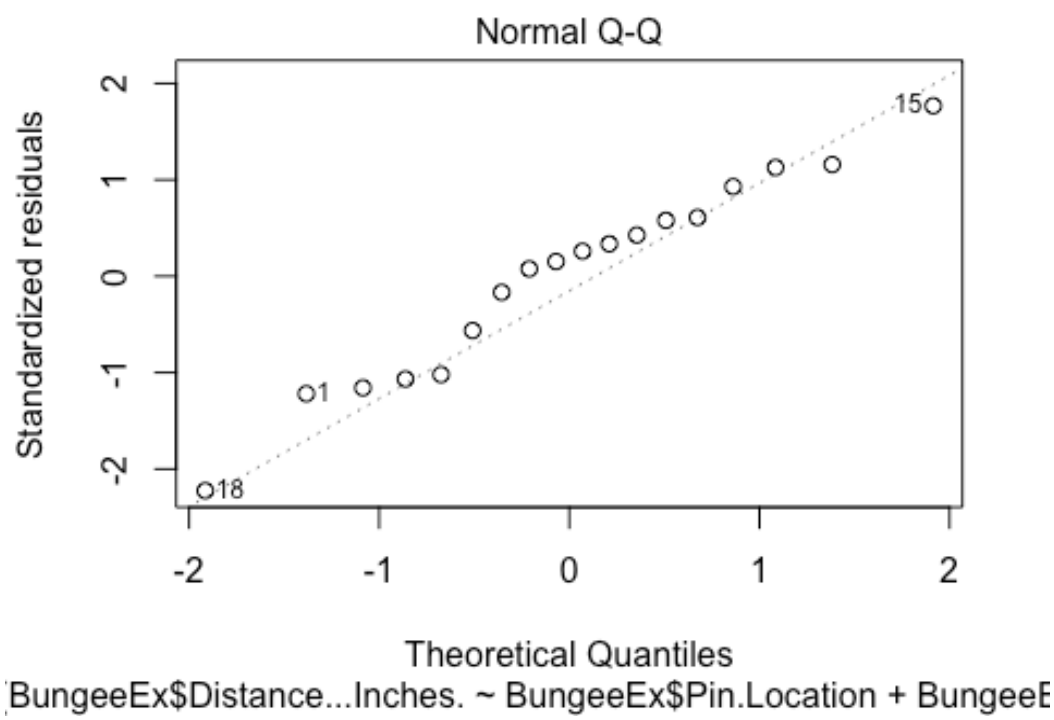
Pin.Location: 0.023703 **<0.05**

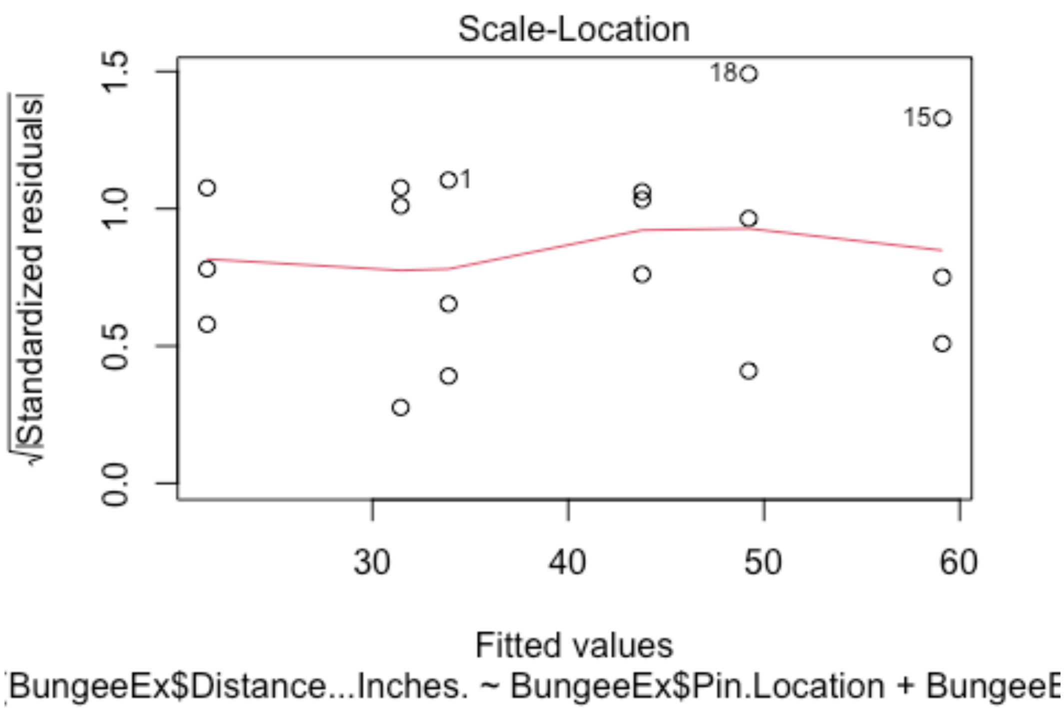
Angle: 0.000187 **<0.05**

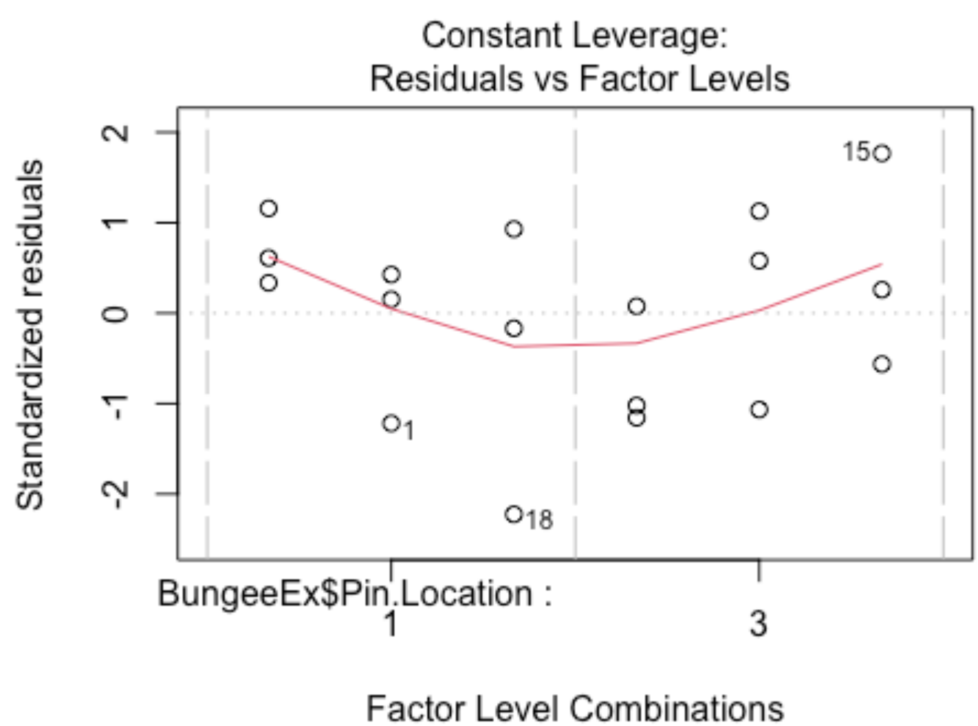
ANOVA Test Plots and Interaction Plot




Residuals vs Fitted
Residuals
Fitted values
BungeeEx$Distance...Inches. ~ BungeeEx$Pin.Location + BungeeE
15
1
18
-15
-5
0
5
10
30
40
50
60


Normal Q-Q
Standardized residuals
Theoretical Quantiles
BungeeEx$Distance...Inches. ~ BungeeEx$Pin.Location + BungeeE
15
1
18
-2
-1
0
1
2


Scale-Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
30
40
50
60
18
15
1
Fitted values
BungeeEx$Distance...Inches. ~ BungeeEx$Pin.Location + BungeeE


Constant Leverage:
Residuals vs Factor Levels
Standardized residuals
-2
-1
0
1
2
15
1
18
BungeeEx$Pin.Location :
1
3
Factor Level Combinations

Conclusion

There seems to be nothing unusaual about the plots. the data seems to follow a straight line on the normal probability plot with 2 extreme outliers on the tail ends of the data distribution. Other than that, everything is fairly normal.

We concluded that the pin location and angle have an effect on the shooting distance of the ball.

2^4 Factorial Design Experiment

Data Collection Layout

For 2^4 factorial design, we used design.ab to generate one replication of a run order for our 2^4 factorial design

```
##    plots r A B C D
## 1    101 1 1 1 2 1
## 2    102 1 1 1 2 2
## 3    103 1 1 2 1 2
## 4    104 1 2 2 2 1
## 5    105 1 2 1 2 1
## 6    106 1 2 2 2 2
## 7    107 1 1 2 2 2
## 8    108 1 1 2 2 1
## 9    109 1 1 2 1 1
## 10   110 1 2 2 1 2
## 11   111 1 2 1 1 2
## 12   112 1 2 1 2 2
## 13   113 1 2 2 1 1
## 14   114 1 1 1 1 2
## 15   115 1 1 1 1 1
## 16   116 1 2 1 1 1
```

Experiment Data and Data Frame

For each of our 4 factors, we had two levels for each factors. They were classified as -1(low) and a +1(high). The different factor levels,and the assigned variables.

Factors and Low and High Levels

	Factor	Low Level(-1)	High Level(+1)
A	Pin Location	Postion 1	Postion 3
B	Bungee Position	Position 2	Position 3
C	Release Angle	140 degrees	170 degrees
D	Ball Type	Yellow	Red

Here is our data that we collected from the experiment.

```
##    Pin_Elevation Bungee_Position Release_Angle Ball_Type response
## 1             -1              -1             1        -1       36
## 2             -1              -1             1         1       35
## 3             -1               1            -1         1       34
## 4              1               1             1        -1       60
## 5              1              -1             1        -1       68
## 6              1               1             1         1       60
## 7             -1               1             1         1       37
## 8             -1               1             1        -1       38
## 9             -1               1            -1        -1       33
## 10             1               1            -1         1       41
## 11             1              -1            -1         1       42
## 12             1              -1             1         1       52
## 13             1               1            -1        -1       51
## 14            -1              -1            -1         1       34
## 15            -1              -1            -1        -1       26
## 16             1              -1            -1        -1       47
```

Null and Alternative Hypothesis Testing

Here are the Hypothesis tests that we used in the experiment. We started at the highest order hypothesis test, which was α_i*β_j hypothesis test.

H$_o$: $\alpha_i = 0$ - Null Hypothesis

H$_a$: $\alpha_i \neq 0$ - Alternative Hypothesis

H$_o$: $\beta_j = 0$ - Null Hypothesis

H$_a$: $\beta_j \neq 0$ - Alternative Hypothesis

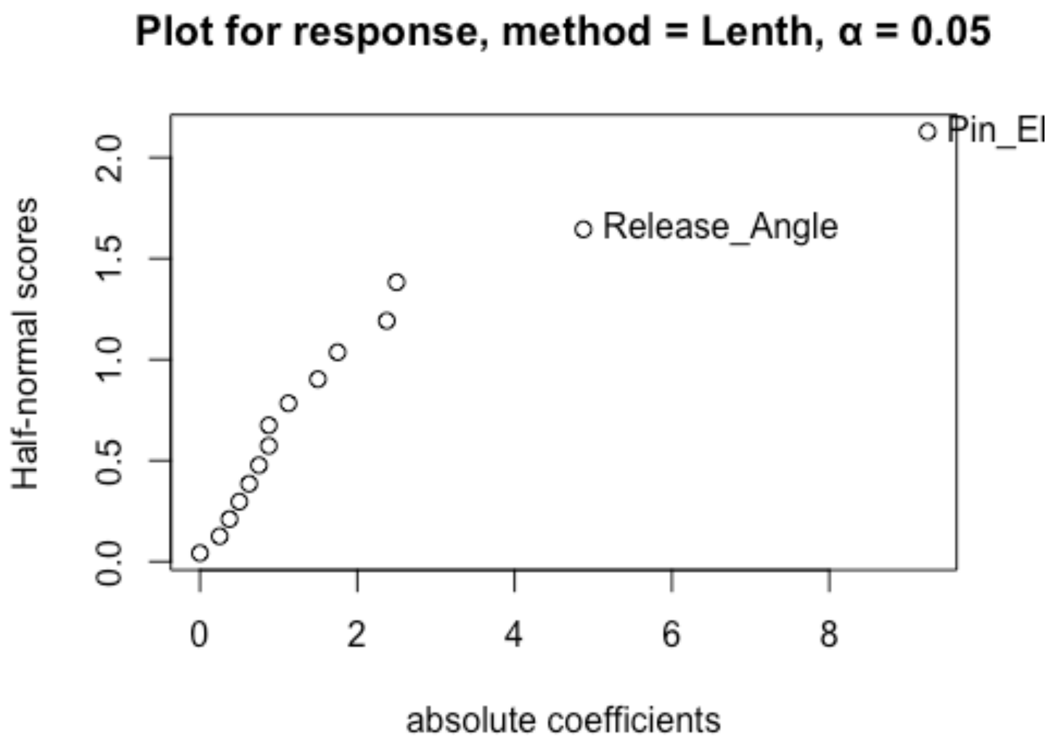
H$_o$: $\alpha\beta_{ij} = 0$ - Null Hypothesis

H$_a$: $\alpha\beta_{ij} \neq 0$ - Alternative Hypothesis

Half Normal Plot

```
##
## Significant effects (alpha=0.05, Lenth method):

## [1] Pin_Elevation Release_Angle
```



From the plot, factors Pin Elevation and Release Angle are significant model terms.

Model Equation

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \epsilon_i$$

$$Distance = 43.37 - 9.25x_{i1} + 4.875x_{i2}$$

ANOVA Model

After running the half normal plot , we concluded that Release Angle and Pin Elevation were significant factors. We run the ANOVA model with those factors and generated the following table.

```
##                Df Sum Sq Mean Sq F value   Pr(>F)    
## Pin_Elevation   1 1369.0  1369.0   51.96 6.86e-06 ***
## Release_Angle   1  380.2   380.2   14.43  0.00221 ** 
## Residuals      13  342.5    26.3                     
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \epsilon_i$$

$$Distance = 29.25 - 18.50x_{i1} + 9.75x_{i2}$$

These are model equations wth their respective coeffents.

Conclusion

From the result, values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case Pin Elevation and Release Angle are significant model terms.

Code

```
### Part 1

library(pwr)
pwr.anova.test(k=3,n=NULL,f=((.5*sqrt((3^2)-1))/(2*3)),sig.level=0.05,power=.
75)
library(agricolae)

treatments<-c("green","yellow","blue")
design<-design.crd(trt=treatments,r=13,seed = 12345)
design$book

library(knitr)

F_levels <- cbind(z$plots,z$r,z$treatments)
kable(F_levels,caption = "Completely Randomized Design ", col.names = c("Plot
","Replication","Color of Ball"))

z <- read.csv("https://raw.githubusercontent.com/Rusty1299/Projects/main/Part
%202%20data%20redoe.csv")
z$treatments <- as.factor(z$treatments)

boxplot(z$distance~z$treatments, col= c("Red","Green","Yellow"), main = "Dist
ance of each ball", xlab = "Treatment balls", ylab = "Distance in inches")
qqnorm(z$distance)
a <- aov(data = z , distance~treatments)
summary(a)

## Part 2
trts<-c(2,3)
design<-design.ab(trt=trts, r=3, design="crd",seed=878900)
design$book

BungeeEx<-read.csv("https://raw.githubusercontent.com/Rusty1299/Projects/main
/Factorial%20Design%20Project.csv")
library(GAD)
BungeeEx$Pin.Location<-as.fixed(BungeeEx$Pin.Location)
BungeeEx$Angle<-as.random(BungeeEx$Angle)

model<-aov(BungeeEx$Distance...Inches.~BungeeEx$Pin.Location*BungeeEx$Angle)
gad(model)
```

```r
model<-aov(BungeeEx$Distance...Inches.~BungeeEx$Pin.Location+BungeeEx$Angle)
gad(model)

interaction.plot(BungeeEx$Angle,BungeeEx$Pin.Location,BungeeEx$Distance...Inc
hes., type = "l", col = 5:7 ,main ="Interraction Plot", ylab = "Distance", xl
ab = "Release Angles", trace.label = "Pin Elevation", lwd = 3, lty = 1)

plot(model)
boxplot(BungeeEx$Distance...Inches.~BungeeEx$Angle, col = 6:9:3, main = "Boxp
lot for Relaease Angle", xlab = "Release Angle", ylab = "Distance")
boxplot(BungeeEx$Distance...Inches.~BungeeEx$Pin.Location, col = 2:4, main =
"Boxplot for Pin Elevation", xlab = "Pin Elevation", ylab = "Distance")

## Part 3

library(agricolae)
#?design.ab
trts<-c(2,2,2,2)
design<-design.ab(trt=trts, r=1, design="crd",seed=878900)

design$book

library(knitr)
A <- c("Pin Location","Postion 1","Postion 3")
B <-c("Bungee Position" ,"Position 2",  "Position 3")
C<-c("Release Angle",   "140 degrees",  "170 degrees")
D<-c("Ball Type",   "Yellow",   "Red")
F_levels <- rbind(A,B,C,D)
colnames(F_levels)<- c("Factor","Low Level(-1)","High Level(+1)")
kable(F_levels,caption = "Factors and Low and High Levels")

library(DoE.base)
Pin_Elevation<-c(-1,-1,-1,1,1,1,-1,-1,-1,1,1,1,1,-1,-1,1)
Bungee_Position<-c(-1,-1,1,1,-1,1,1,1,1,1,-1,-1,1,-1,-1,-1)
Release_Angle<-c(1,1,-1,1,1,1,1,1,-1,-1,-1,1,-1,-1,-1,-1)
Ball_Type<-c(-1,1,1,-1,-1,1,1,-1,-1,1,1,1,-1,1,-1,-1)
response<-c(36,35,34,60,68,60,37,38,33,41,42,52,51,34,26,47)
dat<-data.frame(Pin_Elevation,Bungee_Position,Release_Angle,Ball_Type,respons
e)
dat

model<-lm(response~Pin_Elevation*Bungee_Position*Release_Angle*Ball_Type, dat
a = dat)
#summary(model)
coef(model)

halfnormal(model)
Pin_Elevation<-as.factor(Pin_Elevation)
Bungee_Position<-as.factor(Bungee_Position)
```

```
Release_Angle<-as.factor(Release_Angle)
Ball_Type<-as.factor(Ball_Type)

model1<-aov(response~Pin_Elevation+Release_Angle)
summary(model1)
coef(model1)
```