Exercise 5 Single – Factor Experiments

Kyla Ayop and Charlene Garridos

2024-03-09

Contents

1 R Library

```r
library(FielDHub)
```

```
## Warning: package 'FielDHub' was built under R version 4.3.3
```

```r
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
library(ggplot2)
library(asbio)
```

```
## Warning: package 'asbio' was built under R version 4.3.3

## Loading required package: tcltk
```

```r
library(agricolae)
library(ExpDes)
```

```
##
## Attaching package: 'ExpDes'

## The following objects are masked from 'package:agricolae':
##
##     lastC, order.group, tapply.stat
```

2 Problem

An industrial engineer is conducting an experiment on eye focus time. He is interested in the effect of the distance of the object from the eye on the focus time. Four different distances are of interest. He has five subjects available for the experiment. Because there may be differences among individuals, he decides to conduct the experiment in a randomized block design. The data obtained were as follows.

Distance (Ft)	Subject				
	1	2	3	4	5
1	10	6	6	6	6

Distance (Ft)	Subject				
2	8	6	6	1	6
3	5	3	3	2	2
4	7	4	4	2	3

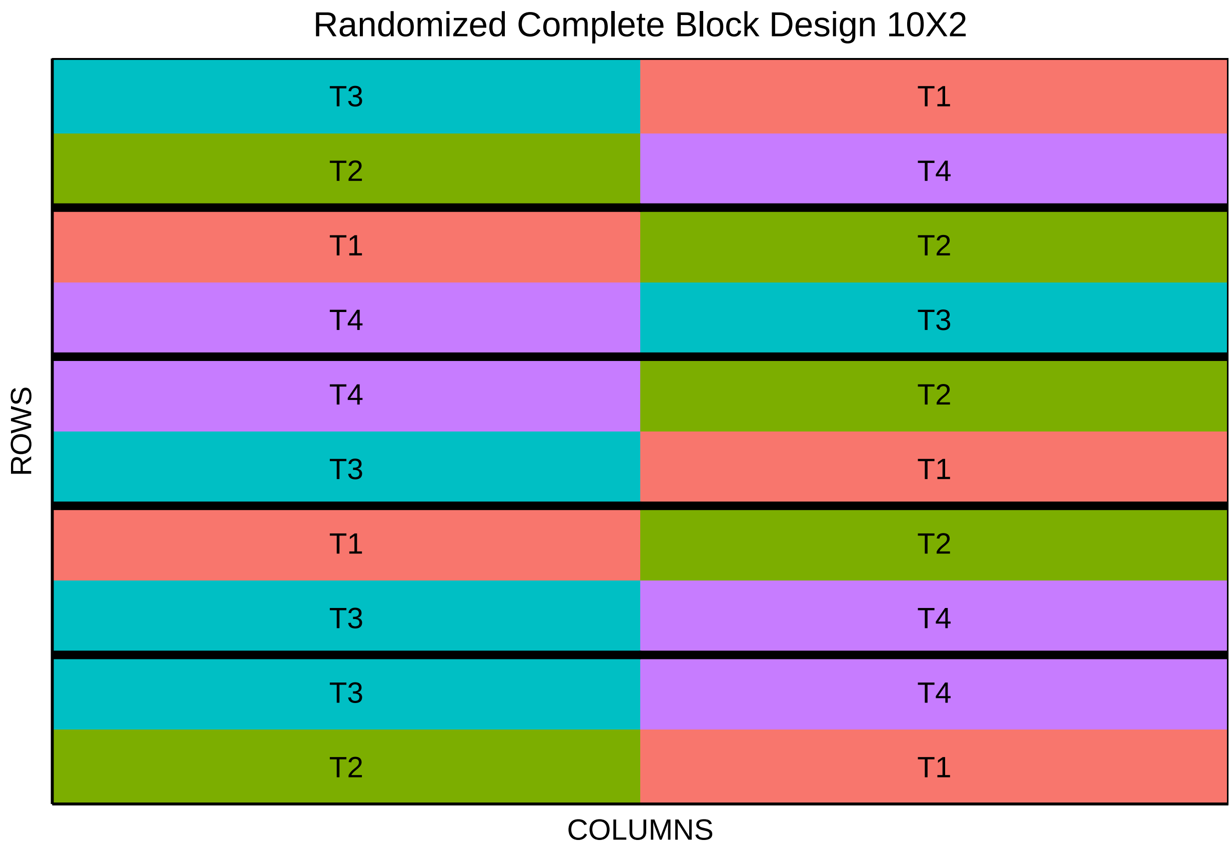
3 Data

```r
EyeFocusTime_rcbd <- RCBD(
 t = 4, # number of treatments
 reps = 5, # number of replications per treatment
 l = 1, # number of locations
 plotNumber = 101, # starting plot number
 locationNames = "IIT", # optional name for each location
 seed = 1237 # seed number to replicate identical randomization
)
print(EyeFocusTime_rcbd)
```

```
## Randomized Complete Block Design (RCBD):
##
## Information on the design parameters:
## List of 7
##  $ blocks              : num 5
##  $ number.of.treatments: num 4
##  $ treatments          : chr [1:4] "T1" "T2" "T3" "T4"
##  $ locations           : num 1
##  $ plotNumber          : num [1:5] 101 201 301 401 501
##  $ locationNames       : chr "IIT"
##  $ seed                : num 1237
##
##  10 First observations of the data frame with the RCBD field book:
##    ID LOCATION PLOT REP TREATMENT
## 1   1      IIT  101   1        T2
## 2   2      IIT  102   1        T1
## 3   3      IIT  103   1        T3
## 4   4      IIT  104   1        T4
## 8   5      IIT  201   2        T3
## 7   6      IIT  202   2        T4
## 6   7      IIT  203   2        T1
## 5   8      IIT  204   2        T2
## 9   9      IIT  301   3        T3
## 10 10      IIT  302   3        T1
```

3.1 RCBD Visualization

```r
plot(EyeFocusTime_rcbd)
```




```r
distance=rep(c(4,6,8,10), each=5)
subject=rep(c(1,2,3,4,5), times =4)
eye_focustime=c(10, 6, 6, 6, 6, 8, 6, 6, 1, 6, 5, 3, 3, 2, 2, 7, 4, 4, 2, 3)
block.subj=as.factor(subject)
trt.distance=as.factor(distance)
eyefocustime_data=data.frame(trt.distance, block.subj, eye_focustime)
eyefocustime_data
```

```
##    trt.distance block.subj eye_focustime
## 1             4          1            10
## 2             4          2             6
## 3             4          3             6
## 4             4          4             6
## 5             4          5             6
## 6             6          1             8
## 7             6          2             6
## 8             6          3             6
## 9             6          4             1
## 10            6          5             6
## 11            8          1             5
```

```
## 12            8          2             3
## 13            8          3             3
## 14            8          4             2
## 15            8          5             2
## 16           10          1             7
## 17           10          2             4
## 18           10          3             4
## 19           10          4             2
## 20           10          5             3
```

```r
str(eyefocustime_data)
```

```
## 'data.frame':    20 obs. of  3 variables:
##  $ trt.distance : Factor w/ 4 levels "4","6","8","10": 1 1 1 1 1 2 2 2 2 2 ...
##  $ block.subj   : Factor w/ 5 levels "1","2","3","4",..: 1 2 3 4 5 1 2 3 4 5 ...
##  $ eye_focustime: num  10 6 6 6 6 8 6 6 1 6 ...
```

3.2 Data Summary

```r
eyefocustime_data %>%
 group_by(trt.distance) %>%
 summarise(
   count = n(),
   mean_eyefocustime = mean(eye_focustime),
   total=sum(eye_focustime),
   sd_eyefocustime=sd(eye_focustime)
 )
```

```
## # A tibble: 4 x 5
##   trt.distance count mean_eyefocustime total sd_eyefocustime
##   <fct>        <int>             <dbl> <dbl>           <dbl>
## 1 4                5               6.8    34            1.79
## 2 6                5               5.4    27            2.61
## 3 8                5               3      15            1.22
## 4 10               5               4      20            1.87
```

```r
eyefocustime_data %>%
 group_by(block.subj) %>%
 summarise(
   count = n(),
   mean_eyefocustime = mean(eye_focustime),
   total=sum(eye_focustime),
   sd_eyefocustime=sd(eye_focustime)
 )
```

```
## # A tibble: 5 x 5
##   block.subj count mean_eyefocustime total sd_eyefocustime
##   <fct>      <int>             <dbl> <dbl>           <dbl>
## 1 1              4              7.5     30            2.08
## 2 2              4              4.75    19            1.5
```

```
## 3 3                4               4.75     19              1.5
## 4 4                4               2.75     11              2.22
## 5 5                4               4.25     17              2.06
```

3.2.1 Visual Inspection

```r
# visual inspection using box plot
ggplot(eyefocustime_data, aes(x = trt.distance, y = eye_focustime)) +
  geom_boxplot(aes(fill = trt.distance),show.legend = FALSE) +
  stat_summary(fun = "mean", geom = "point", shape = 8, size = 2, color = "white") +
  theme_light()
```

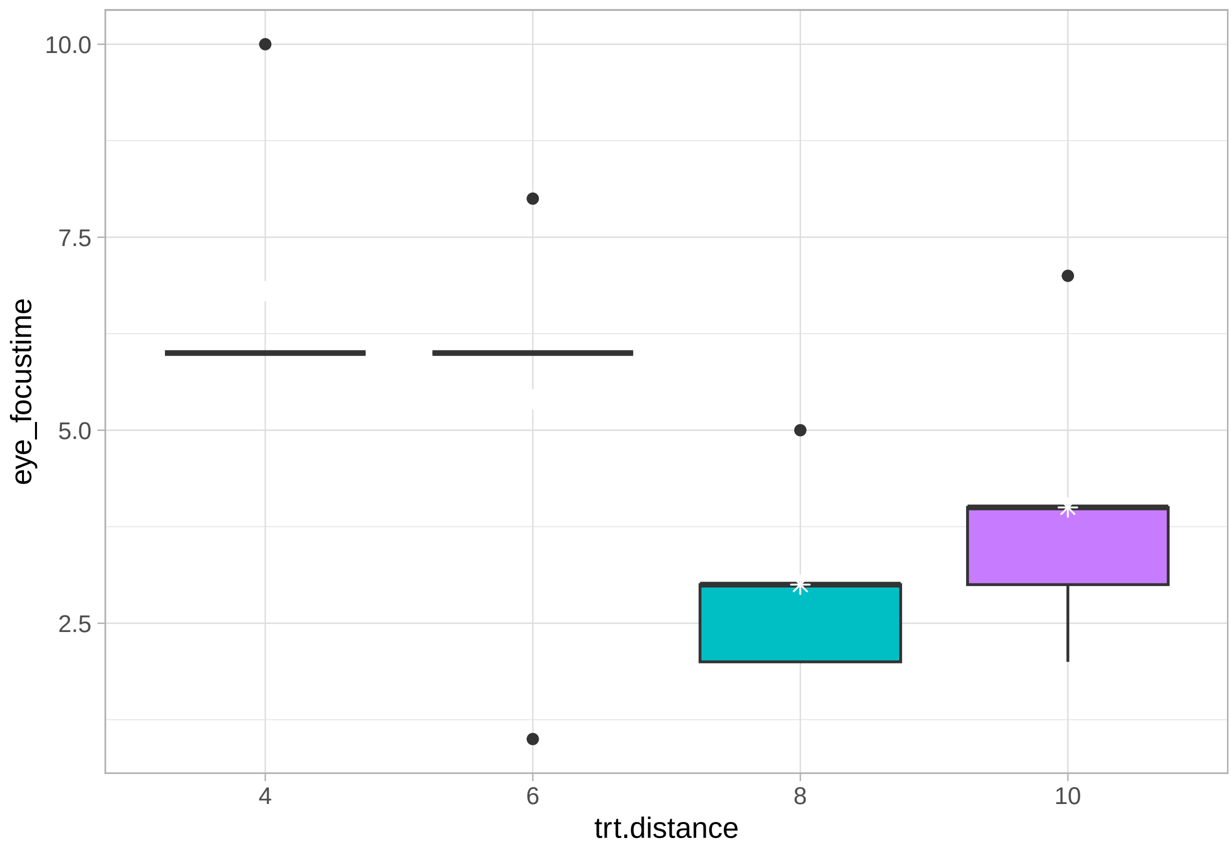


Figure 1: The Box Plot of Treatment Distance

Interpretation

The Figure 1 shows that the distances of 8ft and 10ft are skewed to the left. Additionally, the 10ft distance has a whisker indicating the distribution of data within the population, which means it has a wide range of samples. Meanwhile, the 4ft and 6ft distances show no spread of data and only a horizontal line indicating the median. This implies that the dataset has very low variability and is homogeneous. As for the outliers, it is visible in each distance category (4ft, 6ft, 8ft, and 10ft).

```r
ggplot(eyefocustime_data, aes(x = block.subj, y = eye_focustime)) +
  geom_boxplot(aes( fill = block.subj),show.legend = FALSE) +
  stat_summary(fun = "mean", geom = "point", shape = 8, size = 2, color = "red") +
  theme_light()
```

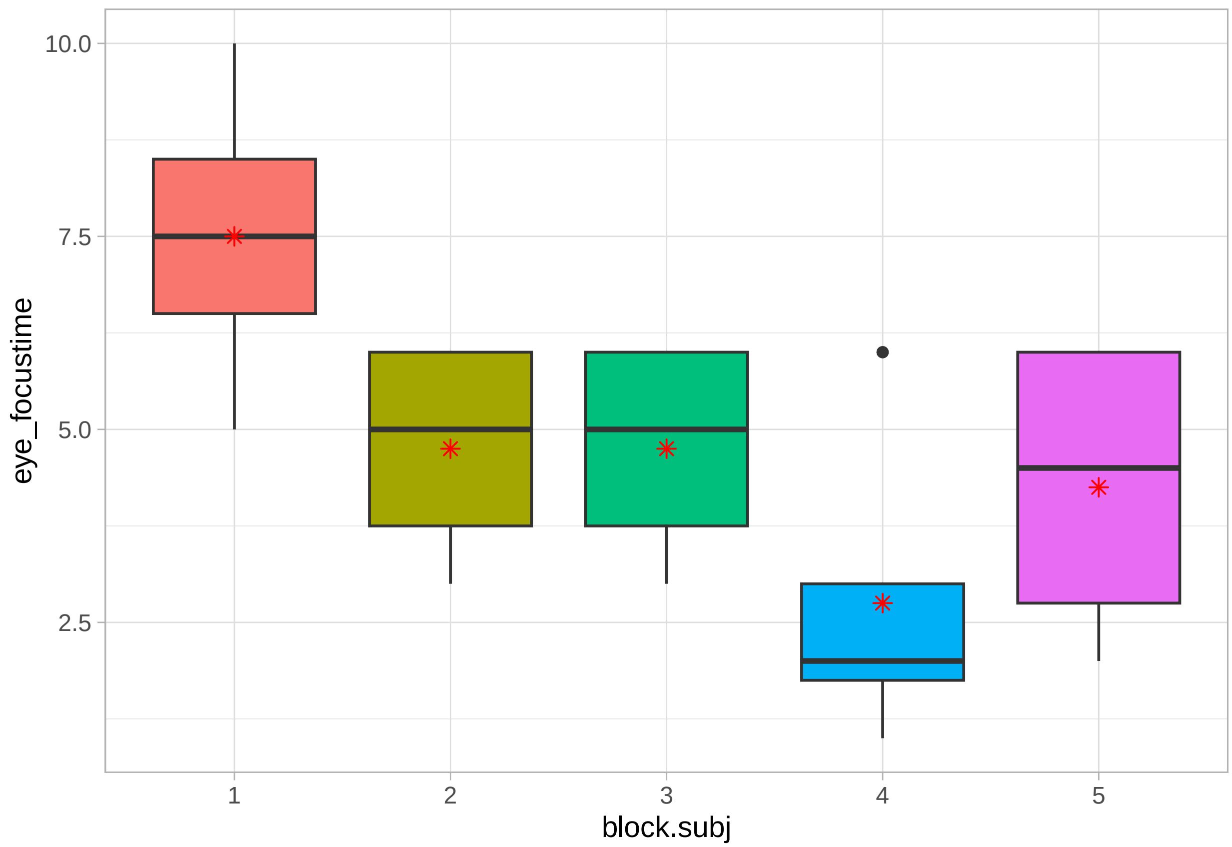


Figure 2: The Box Plot of Block Subject

Interpretation

Figure 2 depicts the five subjects as a blocking factor. Subject 1 is normally distributed with the same value for mean and median. Subjects 2 and 3 are subjectively similar in terms of spread, mean, and median. Subject 4 is skewed to the right, with its mean close to its upper end. Subject 5 is left-skewed, with a mean lower than the median.

4 Analysis of Variance (ANOVA)

```
rcbd_anova=aov(eye_focustime ~ trt.distance + block.subj, eyefocustime_data)
summary(rcbd_anova)
```

```
##              Df Sum Sq Mean Sq F value   Pr(>F)
## trt.distance  3   41.2  13.733   12.88 0.000464 ***
## block.subj    4   47.2  11.800   11.06 0.000539 ***
## Residuals    12   12.8   1.067
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Discussion

Test Statistic: One-way analysis of Variance at $\alpha = 0.05$.

I. Hypothesis Testing for Treatment Effects

H_O: There is no significant difference in the mean focus time among the four different distance of the object from the eye.

$$[\mu_1 = \mu_2 = \mu_3 = \mu_4]$$

H_a: At least one distance gave a different mean focus time among the four different distance of the object from the eye.

$$[\mu_i \neq \mu_j, i \neq j]$$

Decision Rule: Reject H_o if $F_O > F_{[a-1,(a-1)(b-1)]} = 3.49$. Otherwise, do not reject H_o.

Decision: Since $F_O = 12.88 > F_{[a-1,(a-1)(b-1)]} = 3.49$, then H_o is rejected.

Conclusion: At $\alpha = 0.05$, the null hypothesis is rejected. Thus, at least one distance gave a different mean focus time among the four different distance of the object from the eye.

II. Hypothesis Testing for Blocking

H_O: Subject as a blocking factor is not effective.

H_a: Subject as a blocking factor is effective.

Decision Rule: Reject H_o if $F_B > F_{[a-1,(a-1)(b-1)]} = 3.26$. Otherwise, do not reject H_o.

Decision: Since $F_B = 11.132 > F_{[a-1,(a-1)(b-1)]} = 3.26$, then H_o is rejected.

Conclusion: At $\alpha = 0.05$, the null hypothesis is rejected. Thus, blocking the units according to subject is an effective strategy.

5 Test Assumption

5.1 Homogeneity of variances

```
plot(rcbd_anova,which=1, add.smooth = FALSE)
```

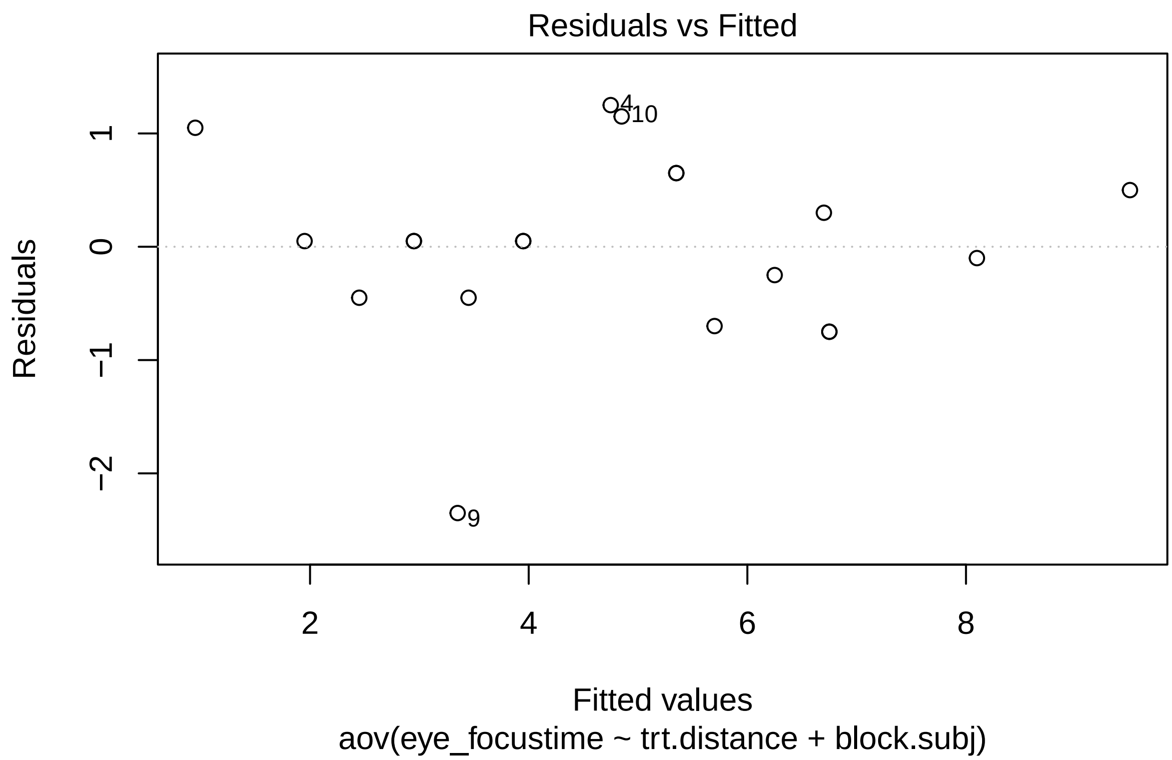


Figure 3: The Residual Vs Fitted Values

Interpretation

On closer examination, Figure 3 subjectively shows no unusual structure like a funnel shape. There is also one outlier; however, it's possible that it implies constant variance (homoscedasticity). This indicates that the variability of the residuals remains relatively consistent across all levels of the predictor variable and doesn't need applying data transformation.

5.2 Normality Assumption

```
plot(rcbd_anova,which=2)
```

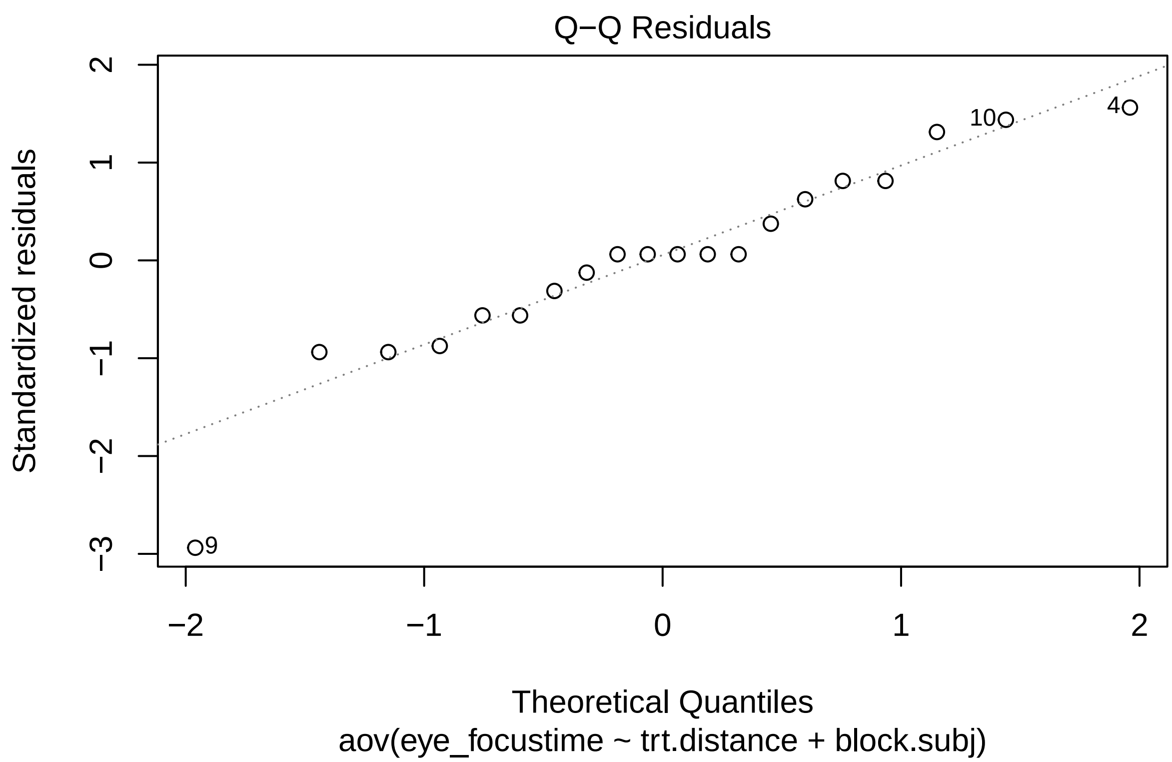


Figure 4: The Normal Q-Q Plot of Residuals

Interpretation

Hypotheses:

$H_o : X_i's$ comes from a normal distribution.

$H_a : X_i's$ do not come from a normal distribution.

The residuals plotted in Figure 4 on the normal probability plot do not fall along a straight line but instead show a wavy pattern. The error distribution may be slightly skewed, with the right tail being longer than the left. This suggests that the normality assumption is not valid, and a data transformation is recommended. Furthermore, there is an indication of an outlier.

5.3 Additivity

```
with(rcbd_anova, interaction.plot(trt.distance, block.subj, eye_focustime,
                                  col=c("blue","red","green")))
```

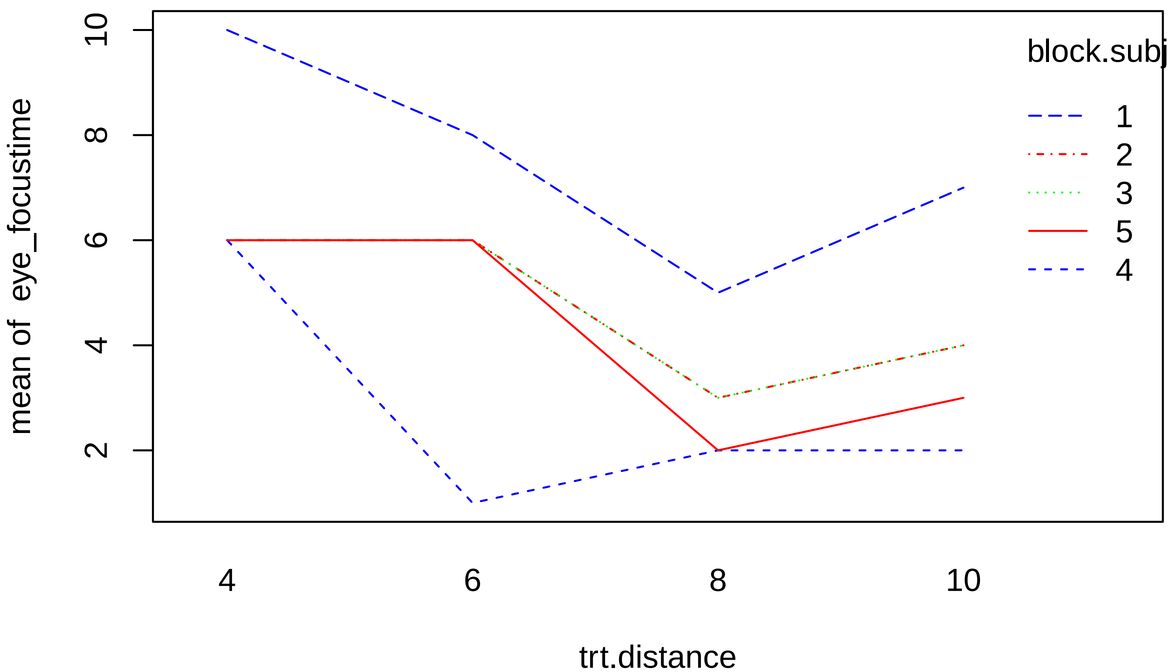


Figure 5: Interaction Plot of Eye Focus Time by Treatment in a Randomized Complete Block Design

Interpretation

Figure 5 shows an Interaction Plot of Eye Focus Time by Treatment in a Randomized Complete Block Design. This interaction plot illustrates how the relationship between treatment distance and eye focus time varies across different subjects, with the subject serving as the blocking factor. The interaction effect suggests that the impact of treatment distance on eye focus time is influenced by the subject. For example, at a distance of 6ft, subjects 2, 3, and 5 exhibit similar mean eye focus times, while subject 4 shows the lowest mean and subject 1 shows the highest mean eye focus time.

```r
with(rcbd_anova, interaction.plot(block.subj, trt.distance,eye_focustime,
                                  col=c("blue","red","green","purple","pink")))
```

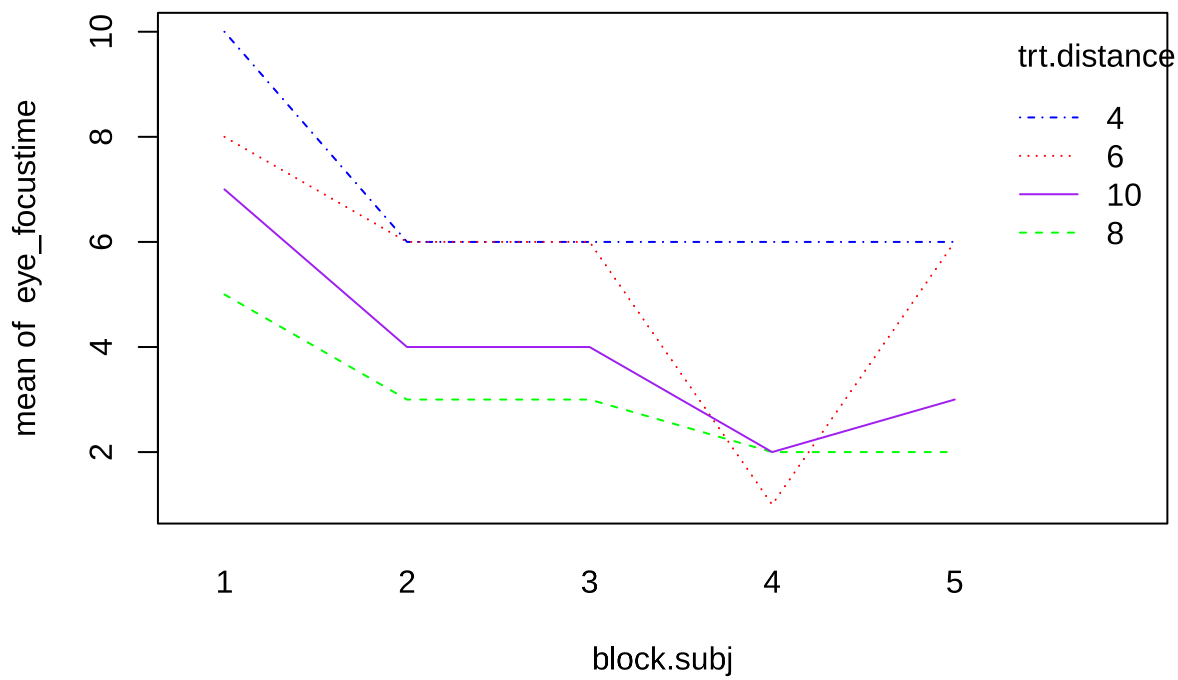


Figure 6: Illustration of Eye Focus Time by Block in a Randomized Complete Block Design

Interpretation

Figure 6 depicts an Interaction Plot of Eye Focus Time by block in a Randomized Complete Block Design. This interaction plot illustrates how the relationship between the blocking factor (subject) and eye focus time varies across different distances, which serve as treatments. The observed interaction effect suggests that the impact of the blocking factor (subject) on eye focus time is influenced by the distance. For example, when considering subject 4 as a block, 6ft is associated with the lowest mean focus time, contrary to 4ft.

```r
#perform Tukey 1df- test for Non-additivity
with(rcbd_anova, tukey.add.test(eye_focustime, trt.distance, block.subj))
```

```
##
## Tukey's one df test for additivity
## F = 0.3731321   Denom df = 11    p-value = 0.5537142
```

Discussion

Hypotheses:

H_o : The effects are additive.

H_a : The effects are not additive.

Decision Rule: Reject H_o if $F_C > F_{\alpha(1,tr-(t+r))}$. Otherwise, do not reject H_o.

Decision: Since the calculated F-value $F_C = 0.3731321 > F_{\alpha(1,tr-(t+r))=3.23}$, H_0 is rejected. Therefore, there is evidence to suggest that the effects are not additive.

Conclusion: At $\alpha = 0.05$, the null hypothesis is rejected. Thus, The effects of different distance of the object from the eye on the focus time are not additive.

6 Multiple Comparison Test

6.1 Duncan Multiple Range Test (DMRT)

```
dmrt=duncan.test(rcbd_anova, "trt.distance", alpha = 0.05)
dmrt
```

```
## $statistics
##    MSerror Df Mean       CV
##   1.066667 12  4.8 21.51657
##
## $parameters
##     test       name.t ntr alpha
##   Duncan trt.distance   4  0.05
##
## $duncan
##      Table CriticalRange
## 2 3.081307      1.423195
## 3 3.225244      1.489676
## 4 3.312453      1.529957
##
## $means
##    eye_focustime      std r        se Min Max Q25 Q50 Q75
## 10           4.0 1.870829 5 0.4618802   2   7   3   4   4
## 4            6.8 1.788854 5 0.4618802   6  10   6   6   6
## 6            5.4 2.607681 5 0.4618802   1   8   6   6   6
## 8            3.0 1.224745 5 0.4618802   2   5   2   3   3
##
## $comparison
## NULL
##
## $groups
##    eye_focustime groups
## 4            6.8      a
## 6            5.4     ab
## 10           4.0     bc
## 8            3.0      c
##
## attr(,"class")
## [1] "group"
```

```r
plot(dmrt)
```

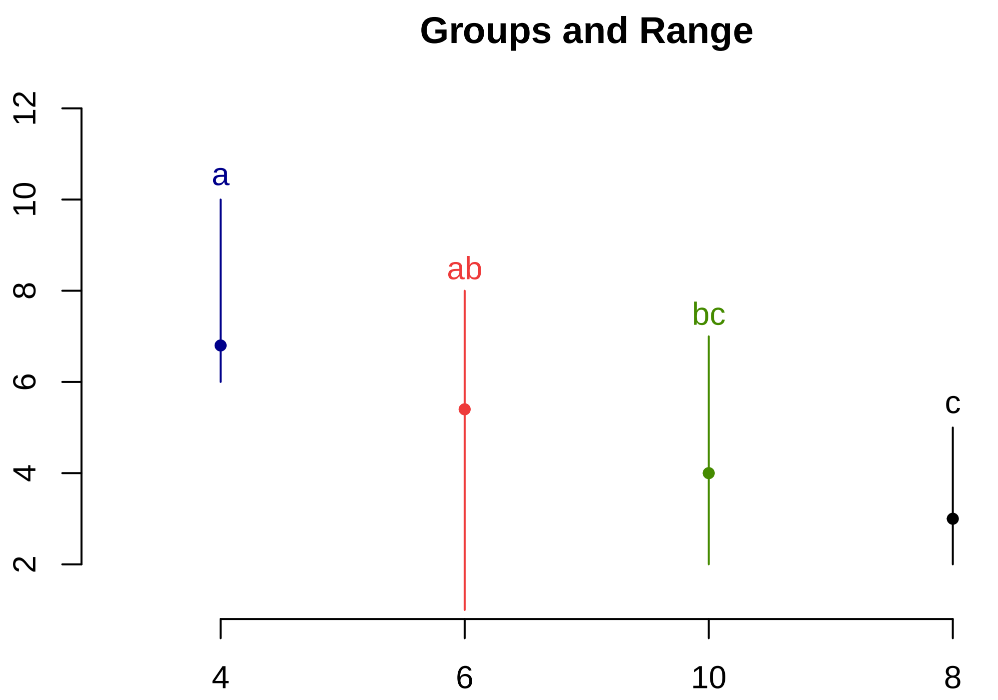


Figure 7: The Plot of Duncan Multiple Range Test (DMRT)

Interpretation

Figure 7 displays the groups and ranges of the four different distance categories. From Figure 7, we can visually determine significant differences among the groups. Specifically, 4ft is not significantly different from 6ft but is significantly different from 10ft and 8ft. Similarly, 6ft is not significantly different from 4ft and 10ft but is significantly different from 8ft. Additionally, 10ft is not significantly different from 8ft but is significantly different from 4ft. Finally, 8ft is not significantly different from 10ft but is significantly different from 4ft and 6ft.

7 Standard Error Computation

```r
MSE = 1.066667
MSTr = 13.733
MSB = 11.800
a = 4
b = 5
```

7.1 Block Mean

$$se(\bar{y}_i) = \sqrt{\frac{MSE}{a}}$$

```
Stand.error_Block.mean <- sqrt(MSE/a)
cat("The standard error of block mean is", Stand.error_Block.mean, "\n")
```

```
## The standard error of block mean is 0.5163979
```

7.2 Treatment Mean

$$se(\bar{y}_i) = \sqrt{\frac{MSE}{b}}$$

```
Stand.error_Treatment.mean <- sqrt(MSE/b)
cat("The standard error of treatment mean is", Stand.error_Treatment.mean, "\n")
```

```
## The standard error of treatment mean is 0.4618803
```

8 Block Efficiency

$$\hat{\sigma}_b^2 = MSE$$
$$\hat{\sigma}_r^2 = \frac{(b-1)MSB + b(a-1)MSE}{ab-1}$$

```
Error.variance_crd = ((b-1)*MSB + b*(a-1)*MSE)/((a*b)-1)
cat("The error variance of CRD is", Error.variance_crd, "\n")
```

```
## The error variance of CRD is 3.326316
```

```
Error.variance_rcbd = MSE
cat("The error variance of RCBD is", Error.variance_rcbd, "\n")
```

```
## The error variance of RCBD is 1.066667
```

$$RE = \frac{(df_b+1)(df_r+3)}{(df_b+3)(df_r+1)} \cdot \frac{\hat{\sigma}_r^2}{\hat{\sigma}_b^2}$$

```
df_b = (a-1)*(b-1)
df_r = (a*b)-a
Relative.efficiency = (((df_b + 1)*(df_r + 3))/((df_b + 3)*(df_r + 1)))*
  (Error.variance_crd/Error.variance_rcbd)

cat("Hence, the relative efficiency of RCDB to CRD is estimated to be",
    Relative.efficiency,"\n")
```

```
## Hence, the relative efficiency of RCDB to CRD is estimated to be 3.020588
```

Therefore, CRD will need about 3 times as many replications to obtain the same precision as obtained by blocking on subject types.