Project2

Holly Cui, Elyse McFalls

2023-12-04

In our design, there are 16 conditions of paper planes as follows:

ID	Treatment	ID	Treatment
a	Only high rotor length	bd	High leg length & leg clip
b	Only high leg length	cd	High leg width & leg clip
c	Only high leg width	abc	High rotor length & high leg length & high leg width
d	Only leg clip	abd	High rotor length & high leg length & leg clip
ab	High rotor length & high leg length	acd	High rotor length & high leg width & leg clip
ac	High rotor length & high leg width	bcd	High leg length & high leg width & leg clip
ad	High rotor length & leg clip	abcd	High on all factors
bc	High leg length & high leg width	1	Low on all factors

where the levels of highs and lows are defined as:

- **Rotor length**: low = 7.5 cm, high = 8.5 cm
- **Leg length**: low = 7.5 cm, high = 12.0 cm
- **Leg width**: low = 3.2 cm, high = 5.0 cm
- **Leg clip**: no, yes

```
# randomize treatment for each flight
set.seed(123)
trt = c(rep("a", 5), rep("b", 5), rep("c", 5), rep("d", 5),
        rep("ab", 5), rep("ac", 5), rep("ad", 5), rep("bc", 5),
        rep("bd", 5), rep("cd", 5), rep("abc", 5), rep("abd", 5),
        rep("acd", 5), rep("bcd", 5), rep("abcd", 5), rep("1", 5))
sample(trt)
```

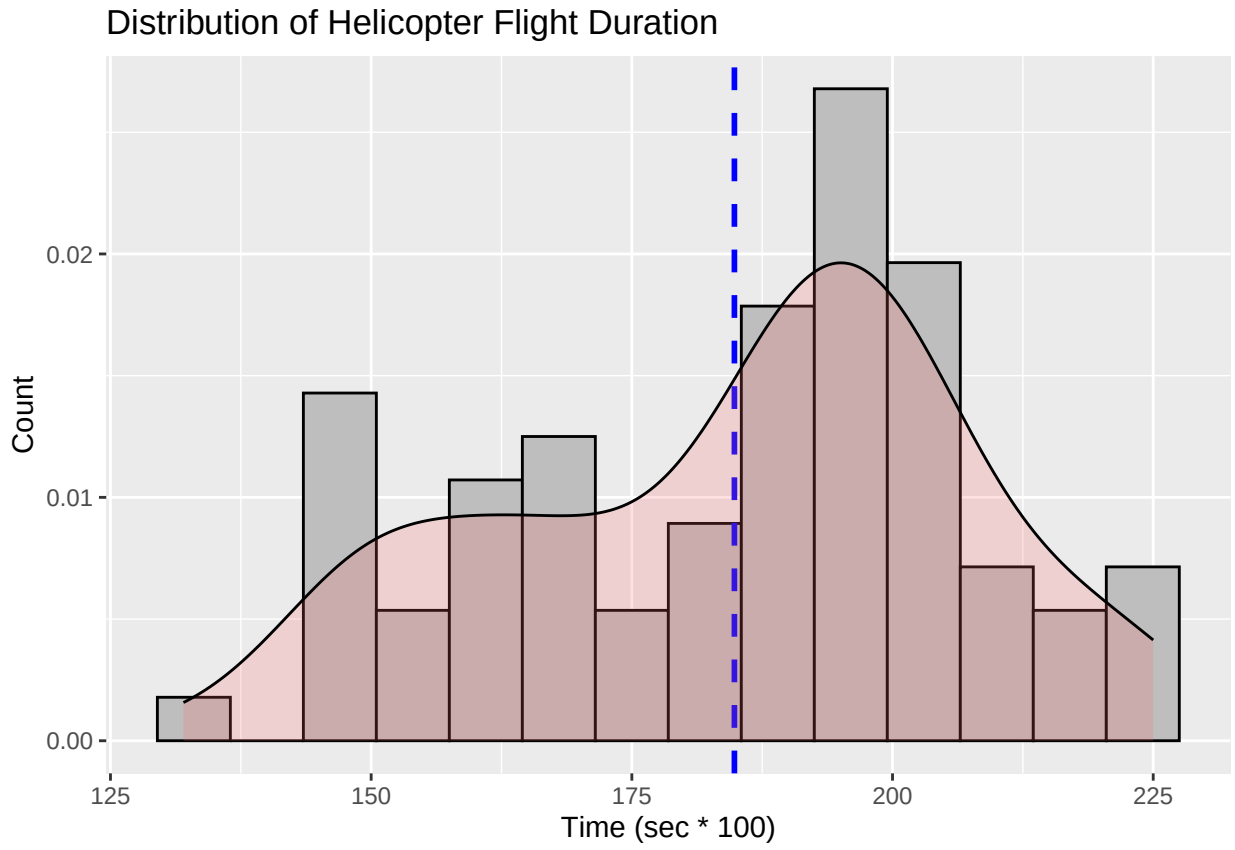
```
##  [1] "ad"   "1"    "abc"  "c"    "bcd"  "bd"   "cd"   "bd"   "1"    "ab"
## [11] "bcd"  "abd"  "b"    "ac"   "b"    "abcd" "bcd"  "acd"  "d"    "bc"
## [21] "abcd" "d"    "abcd" "bc"   "abc"  "c"    "c"    "ad"   "acd"  "bd"
## [31] "bcd"  "acd"  "bd"   "b"    "ab"   "ac"   "abcd" "1"    "abd"  "bc"
## [41] "abcd" "ad"   "ac"   "a"    "b"    "abc"  "c"    "d"    "a"    "bc"
## [51] "abd"  "1"    "b"    "ab"   "abc"  "cd"   "1"    "d"    "d"    "ac"
## [61] "c"    "acd"  "bd"   "cd"   "abc"  "ac"   "a"    "abd"  "abd"  "a"
## [71] "bc"   "ab"   "ad"   "a"    "bcd"  "ab"   "ad"   "cd"   "cd"   "acd"
```

```
plane = read.csv("paperplane.csv", header = T)
dim(plane)
```

```
## [1] 80  6
```

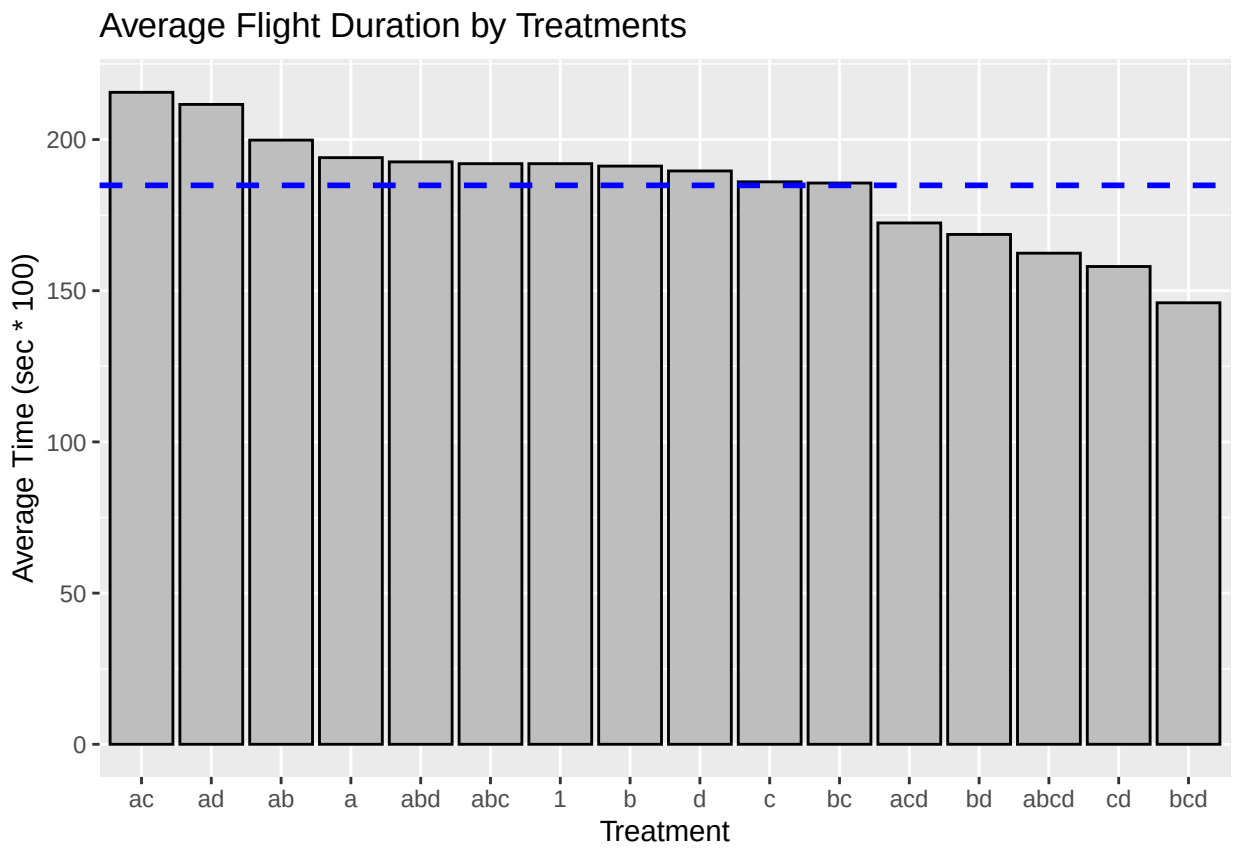
EDA:

```
ggplot(plane, aes(x = time*100)) +
  geom_histogram(aes(y=after_stat(density)),
                 binwidth=7, color="black", fill="grey") +
  geom_vline(aes(xintercept = mean(time)*100),
             color="blue", linetype="dashed", linewidth=1) +
  geom_density(alpha=.2, fill="#FF6666") +
  labs(title = "Distribution of Helicopter Flight Duration",
       x = "Time (sec * 100)", y = "Count")
```



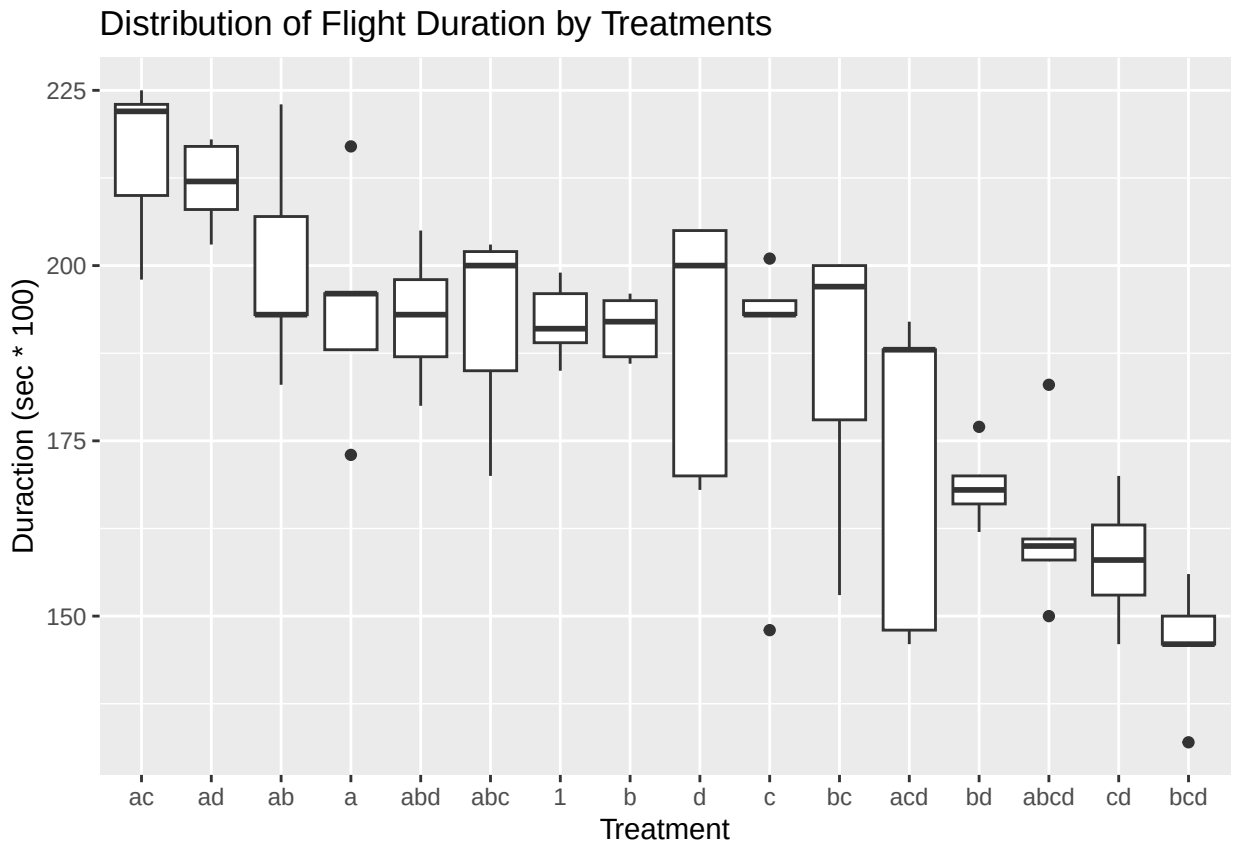
- Insights:

```
plane %>%
  group_by(trt) %>%
  summarise(mean_time = mean(time*100)) %>%
  arrange(desc(mean_time)) %>%
  ggplot(aes(y = mean_time, x = fct_rev(fct_reorder(trt, mean_time)))) +
  geom_col(color="black", fill="grey") +
  geom_hline(aes(yintercept = mean(mean_time)),
             color="blue", linetype="dashed", linewidth=1) +
  labs(title = "Average Flight Duration by Treatments",
       x = "Treatment", y = "Average Time (sec * 100)")
```



- insights: the presence of a - long rotor seems to be helping long duration. the presence of d is discouraging long duration.

```
plane %>%
  mutate(trt = factor(trt, levels = c("ac", "ad", "ab", "a",
                                      "abd", "abc", "1", "b",
                                      "d", "c", "bc", "acd",
                                      "bd", "abcd", "cd", "bcd"))) %>%
ggplot(aes(x = trt, y = time*100)) +
  geom_boxplot() +
  labs(title = "Distribution of Flight Duration by Treatments",
       x = "Treatment", y = "Duraction (sec * 100)")
```




```
tapply(plane[,1], plane[,6], sd)
```

```
##          1          a         ab        abc       abcd        abd         ac
## 0.05567764 0.15921683 0.15530615 0.14300350 0.12300406 0.09659193 0.11458621
##        acd         ad          b         bc        bcd         bd          c
## 0.23255107 0.06268971 0.04549725 0.20403431 0.08831761 0.05549775 0.21494185
##         cd          d
## 0.09192388 0.18928814
```

- Insights: box plots are difficult to interpret with only 5 or so observations per box, but we can still observe
  - an outlier in treatment bcd and other groups: we may want to estimate the model with and without that point to see if our conclusions are overly sensitive to that point. But since most outliers in groups are within the normal range, we don't worry too much.
  - Along w previous plot, we might need to resolve the problem of distribution of time is right-skewed.
  - The var for group d and acd are relatively larger. However, since the total duration is measured under second, the largest difference is only several miliseconds away, so we can use regression under the equal var assumption. (from the math calculation, most of them confirmed to be the same –> with very small sample size, it is difficult to conclude definitively that the variances are different across groups.)

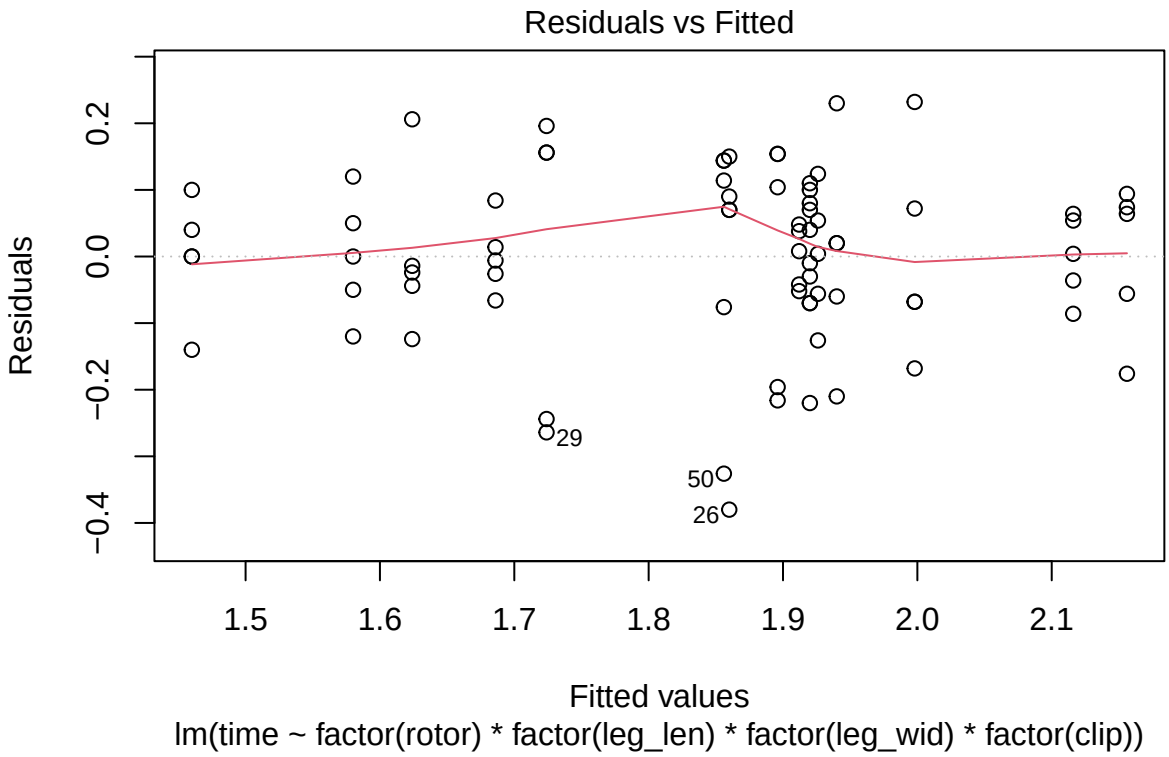
Modeling:

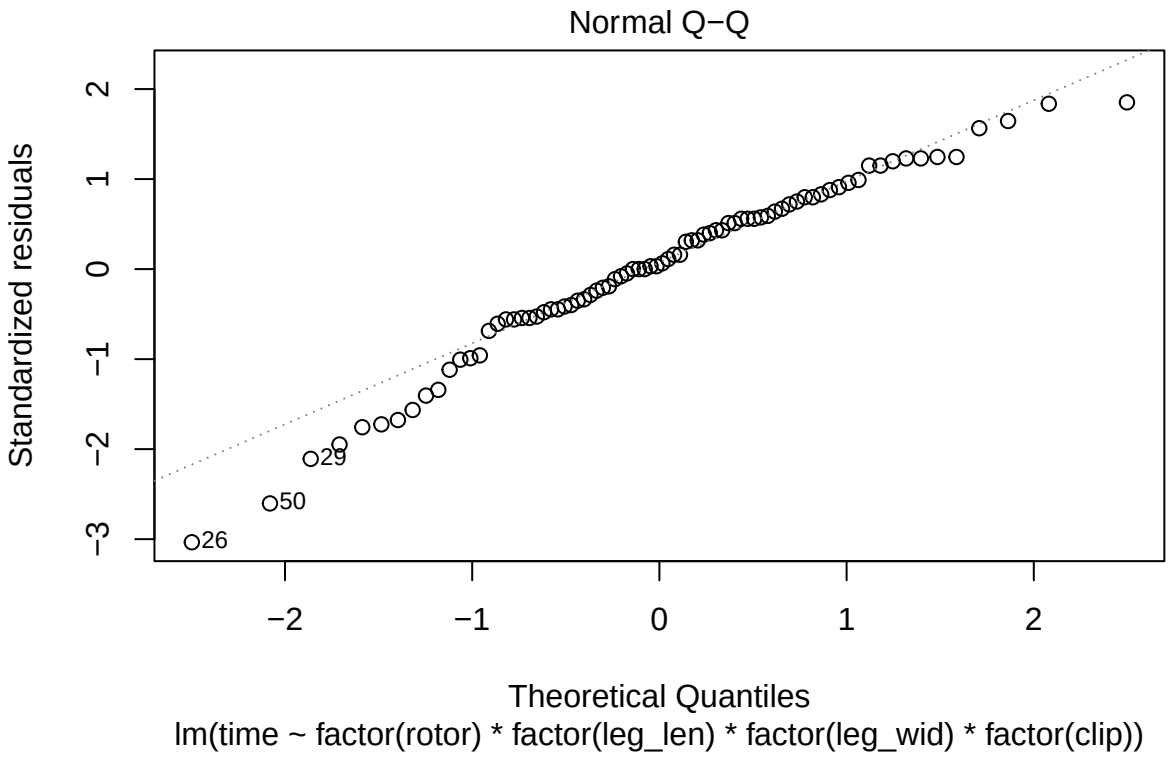
```
reg_inter = lm(time ~ factor(rotor) * factor(leg_len) * factor(leg_wid) * factor(clip), data = plane)
summary(reg_inter)
```

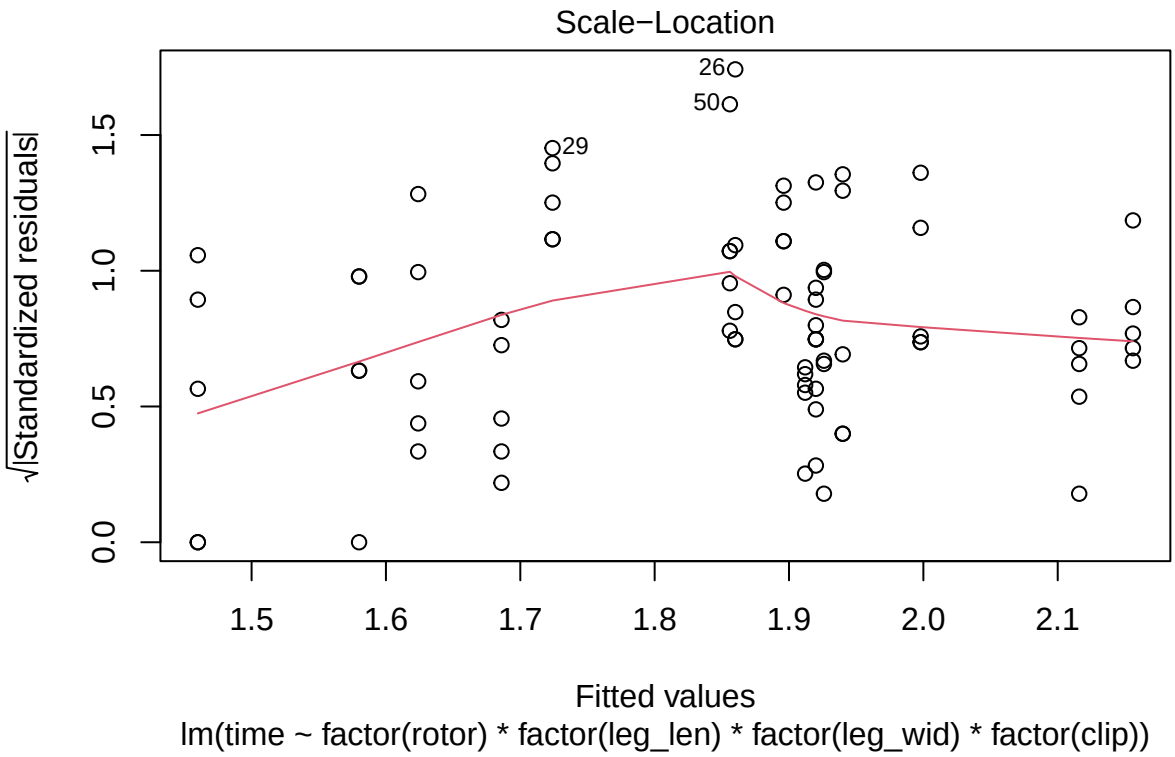
```
##
## Call:
## lm(formula = time ~ factor(rotor) * factor(leg_len) * factor(leg_wid) *
##     factor(clip), data = plane)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3800 -0.0665  0.0060  0.0855  0.2320
##
## Coefficients:
##                                                                  Estimate
## (Intercept)                                                       1.92000
## factor(rotor)1                                                    0.02000
## factor(leg_len)1                                                 -0.00800
## factor(leg_wid)1                                                 -0.06000
## factor(clip)1                                                    -0.02400
## factor(rotor)1:factor(leg_len)1                                   0.06600
## factor(rotor)1:factor(leg_wid)1                                   0.27600
## factor(leg_len)1:factor(leg_wid)1                                 0.00400
## factor(rotor)1:factor(clip)1                                      0.20000
## factor(leg_len)1:factor(clip)1                                   -0.20200
## factor(leg_wid)1:factor(clip)1                                   -0.25600
## factor(rotor)1:factor(leg_len)1:factor(leg_wid)1                 -0.29800
## factor(rotor)1:factor(leg_len)1:factor(clip)1                    -0.04600
## factor(rotor)1:factor(leg_wid)1:factor(clip)1                    -0.35200
## factor(leg_len)1:factor(leg_wid)1:factor(clip)1                   0.08600
## factor(rotor)1:factor(leg_len)1:factor(leg_wid)1:factor(clip)1    0.29800
##                                                                  Std. Error
## (Intercept)                                                         0.06262
## factor(rotor)1                                                      0.08855
## factor(leg_len)1                                                    0.08855
## factor(leg_wid)1                                                    0.08855
## factor(clip)1                                                       0.08855
## factor(rotor)1:factor(leg_len)1                                     0.12524
## factor(rotor)1:factor(leg_wid)1                                     0.12524
## factor(leg_len)1:factor(leg_wid)1                                   0.12524
## factor(rotor)1:factor(clip)1                                        0.12524
## factor(leg_len)1:factor(clip)1                                      0.12524
## factor(leg_wid)1:factor(clip)1                                      0.12524
## factor(rotor)1:factor(leg_len)1:factor(leg_wid)1                    0.17711
## factor(rotor)1:factor(leg_len)1:factor(clip)1                       0.17711
## factor(rotor)1:factor(leg_wid)1:factor(clip)1                       0.17711
## factor(leg_len)1:factor(leg_wid)1:factor(clip)1                     0.17711
## factor(rotor)1:factor(leg_len)1:factor(leg_wid)1:factor(clip)1      0.25047
##                                                                  t value Pr(>|t|)
## (Intercept)                                                       30.662   <2e-16
## factor(rotor)1                                                     0.226   0.8220
## factor(leg_len)1                                                  -0.090   0.9283
## factor(leg_wid)1                                                  -0.678   0.5005
## factor(clip)1                                                     -0.271   0.7872
```

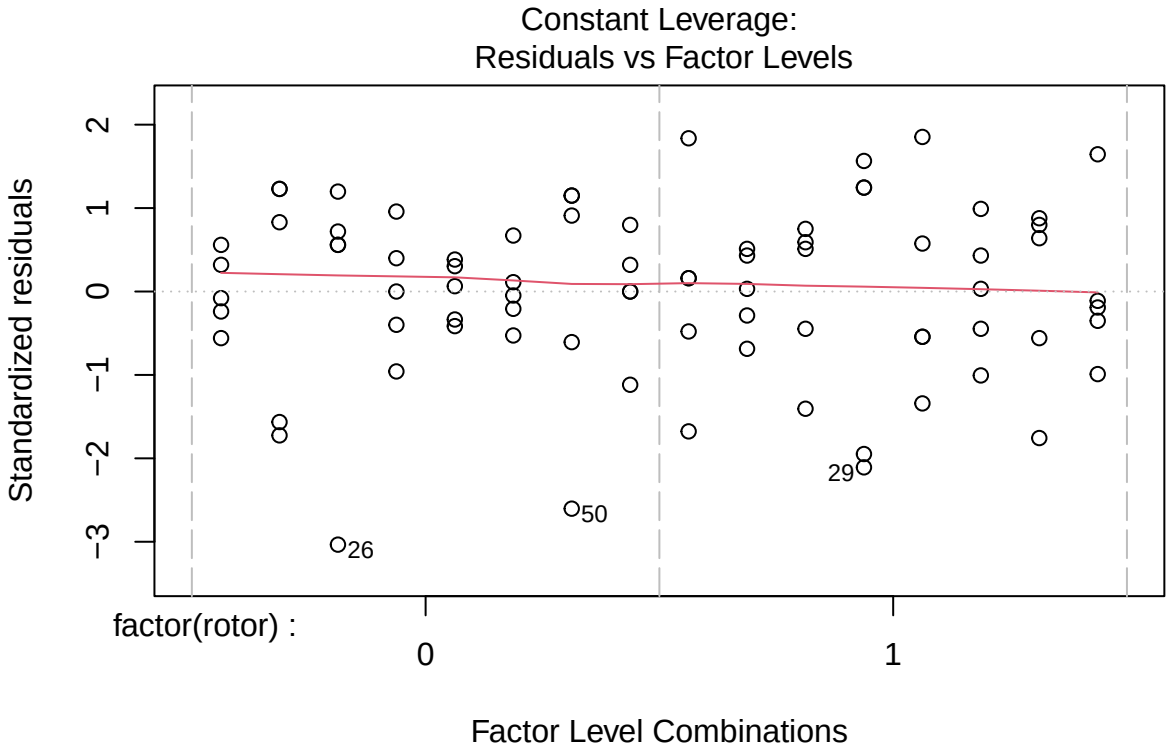
```
## factor(rotor)1:factor(leg_len)1                                     0.527   0.6000
## factor(rotor)1:factor(leg_wid)1                                     2.204   0.0311
## factor(leg_len)1:factor(leg_wid)1                                   0.032   0.9746
## factor(rotor)1:factor(clip)1                                        1.597   0.1152
## factor(leg_len)1:factor(clip)1                                     -1.613   0.1117
## factor(leg_wid)1:factor(clip)1                                     -2.044   0.0451
## factor(rotor)1:factor(leg_len)1:factor(leg_wid)1                   -1.683   0.0973
## factor(rotor)1:factor(leg_len)1:factor(clip)1                      -0.260   0.7959
## factor(rotor)1:factor(leg_wid)1:factor(clip)1                      -1.987   0.0512
## factor(leg_len)1:factor(leg_wid)1:factor(clip)1                     0.486   0.6289
## factor(rotor)1:factor(leg_len)1:factor(leg_wid)1:factor(clip)1      1.190   0.2385
##
## (Intercept)                                                       ***
## factor(rotor)1
## factor(leg_len)1
## factor(leg_wid)1
## factor(clip)1
## factor(rotor)1:factor(leg_len)1
## factor(rotor)1:factor(leg_wid)1                                   *
## factor(leg_len)1:factor(leg_wid)1
## factor(rotor)1:factor(clip)1
## factor(leg_len)1:factor(clip)1
## factor(leg_wid)1:factor(clip)1                                    *
## factor(rotor)1:factor(leg_len)1:factor(leg_wid)1                  .
## factor(rotor)1:factor(leg_len)1:factor(clip)1
## factor(rotor)1:factor(leg_wid)1:factor(clip)1                     .
## factor(leg_len)1:factor(leg_wid)1:factor(clip)1
## factor(rotor)1:factor(leg_len)1:factor(leg_wid)1:factor(clip)1
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.14 on 64 degrees of freedom
## Multiple R-squared:  0.6807, Adjusted R-squared:  0.6058
## F-statistic: 9.094 on 15 and 64 DF,  p-value: 8.164e-11
```

```
plot(reg_inter)
```


Residuals vs Fitted
Residuals
0.2
0.0
−0.2
−0.4
29
50
26
1.5
1.6
1.7
1.8
1.9
2.0
2.1
Fitted values
lm(time ~ factor(rotor) * factor(leg_len) * factor(leg_wid) * factor(clip))


Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(time ~ factor(rotor) * factor(leg_len) * factor(leg_wid) * factor(clip))
29
50
26


Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
26
50
29
1.5
1.6
1.7
1.8
1.9
2.0
2.1
Fitted values
lm(time ~ factor(rotor) * factor(leg_len) * factor(leg_wid) * factor(clip))


Constant Leverage:
Residuals vs Factor Levels
Standardized residuals
2
1
0
−1
−2
−3
26
50
29
factor(rotor) :
0
1
Factor Level Combinations