322 Final Project

Thomas Schaefer & Pouya Mohammadi

4/23/2020

Main Objectives

Determine the answers to the following questions:

1. Which factors seem to be the most important for making helicopters that fly longer?
2. Is there any evidence that the effect of rotor length differs by leg width?
3. What would you recommend as the ideal combination to make the helicopter fly long?

Design

We considered implementing both a half factorial design and a full factorial design for this project. While we knew that we could extract more accurate estimate for the effect of the helicopter design on the flight length by doing a full factorial design, we also realized that a half factorial design would save us time and paper. Therefore, we designed both a half factorial design and a full factorial design, both of which we believed would give us enough information to answer the above objectives. Both of these designs included testing each helicopter variant 4 times and record the design. In our design, a value of 1 for a factor relates to a high assignment and a value of 0 relates to a low assignment for the factor. For the half factorial design, we would only construct helicopters with the following factors:

(rotor length, leg width, paperclip, leg length) (1,1,1,1) (1,0,1,1) (0.1,1,1) (0,0,1,1) (1,1,0,1) (1,1,0,0) (1,1,1,0) (0,0,0,0)

The full factorial design would include the variants above, in addition to the following variants:

(rotor length, leg width, paperclip, leg length) (1,0,0,1) (0,1,0,1) (0,0,0,1) (1,0,1,0) (0,1,1,0) (0,0,1,0) (1,0,0,0) (0,1,0,0)

After beginning to construct and test the helicopters, we realized that we had both the time and resources to conduct a full factorial design. Therefore we conducted a full factorial design on the 4 factors.

Methods

To test the helicopters, we first designed and printed them using graphic design software Inkscape to limit errors. Each helicopter was dropped four times from a height of 2 meters off the ground. The drop time of the helicopters was measure from the point it was dropped to the point of first contact with the ground.

Results

```r
our_data <- read.xlsx("HelicopterData.xlsx", sheetIndex = 1)

our_data
```

```
##    Rotar Leg.Width Leg.Length Paperclip Trial.1 Trial.2 Trial.3 Trial.4
## 1      H         L          L         L    1.84    2.03    1.90    1.96
## 2      H         L          L         H    1.34    1.25    1.16    1.88
## 3      H         L          H         L    1.82    2.06    1.92    1.95
## 4      H         L          H         H    1.45    1.45    1.78    1.72
## 5      H         H          L         L    2.58    1.92    1.89    2.20
## 6      H         H          L         H    2.41    2.55    2.50    2.35
## 7      H         H          H         L    1.92    1.70    2.05    1.90
## 8      H         H          H         H    1.85    1.78    2.13    1.61
## 9      L         L          L         L    1.79    2.05    1.85    1.95
## 10     L         L          L         H    1.19    1.35    1.54    1.48
## 11     L         L          H         L    1.96    1.78    1.76    2.18
## 12     L         L          H         H    1.70    2.01    1.93    1.88
## 13     L         H          L         L    1.92    2.12    1.93    1.85
## 14     L         H          L         H    1.85    1.58    1.88    1.82
## 15     L         H          H         L    2.01    1.94    1.92    1.88
## 16     L         H          H         H    2.02    1.78    1.78    1.82
```

Unfortunately, we did not format the data correctly, so we will fix the formatting to help make our analysis easier.

```r
makeBinary <- function(letter){
  l = as.character(letter)
  if (letter == "H"){
    return(1)
  }
  else if (letter == "L"){
    return(0)
  }
}

our_data$Rotar = mapply(makeBinary, our_data$Rotar)
our_data$Paperclip = mapply(makeBinary, our_data$Paperclip)
our_data$Leg.Width = mapply(makeBinary, our_data$Leg.Width)
our_data$Leg.Length = mapply(makeBinary, our_data$Leg.Length)

drops <- c("Trial.2", "Trial.3", "Trial.4")
trial_1 <- our_data[ , !(names(our_data) %in% drops)]
names(trial_1) <- c("Rotar", "Leg.Width", "Leg.Length", "Paperclip", "y")

drops <- c("Trial.1", "Trial.3", "Trial.4")
trial_2 <- our_data[ , !(names(our_data) %in% drops)]
names(trial_2) <- c("Rotar", "Leg.Width", "Leg.Length", "Paperclip", "y")

drops <- c("Trial.1", "Trial.2", "Trial.4")
trial_3 <- our_data[ , !(names(our_data) %in% drops)]
names(trial_3) <- c("Rotar", "Leg.Width", "Leg.Length", "Paperclip", "y")

drops <- c("Trial.2", "Trial.3", "Trial.1")
trial_4 <- our_data[ , !(names(our_data) %in% drops)]
names(trial_4) <- c("Rotar", "Leg.Width", "Leg.Length", "Paperclip", "y")
```

```r
df <- rbind(trial_1, trial_2)
df <- rbind(df, trial_3)
df <- rbind(df, trial_4)
df <- df %>% arrange(desc(Rotar), desc(Leg.Width), desc(Leg.Length), desc(Paperclip))
df
```

```
##    Rotar Leg.Width Leg.Length Paperclip    y
## 1      1         1          1         1 1.85
## 2      1         1          1         1 1.78
## 3      1         1          1         1 2.13
## 4      1         1          1         1 1.61
## 5      1         1          1         0 1.92
## 6      1         1          1         0 1.70
## 7      1         1          1         0 2.05
## 8      1         1          1         0 1.90
## 9      1         1          0         1 2.41
## 10     1         1          0         1 2.55
## 11     1         1          0         1 2.50
## 12     1         1          0         1 2.35
## 13     1         1          0         0 2.58
## 14     1         1          0         0 1.92
## 15     1         1          0         0 1.89
## 16     1         1          0         0 2.20
## 17     1         0          1         1 1.45
## 18     1         0          1         1 1.45
## 19     1         0          1         1 1.78
## 20     1         0          1         1 1.72
## 21     1         0          1         0 1.82
## 22     1         0          1         0 2.06
## 23     1         0          1         0 1.92
## 24     1         0          1         0 1.95
## 25     1         0          0         1 1.34
## 26     1         0          0         1 1.25
## 27     1         0          0         1 1.16
## 28     1         0          0         1 1.88
## 29     1         0          0         0 1.84
## 30     1         0          0         0 2.03
## 31     1         0          0         0 1.90
## 32     1         0          0         0 1.96
## 33     0         1          1         1 2.02
## 34     0         1          1         1 1.78
## 35     0         1          1         1 1.78
## 36     0         1          1         1 1.82
## 37     0         1          1         0 2.01
## 38     0         1          1         0 1.94
## 39     0         1          1         0 1.92
## 40     0         1          1         0 1.88
## 41     0         1          0         1 1.85
## 42     0         1          0         1 1.58
## 43     0         1          0         1 1.88
## 44     0         1          0         1 1.82
## 45     0         1          0         0 1.92
## 46     0         1          0         0 2.12
```

```
## 47     0          1          0          0 1.93
## 48     0          1          0          0 1.85
## 49     0          0          1          1 1.70
## 50     0          0          1          1 2.01
## 51     0          0          1          1 1.93
## 52     0          0          1          1 1.88
## 53     0          0          1          0 1.96
## 54     0          0          1          0 1.78
## 55     0          0          1          0 1.76
## 56     0          0          1          0 2.18
## 57     0          0          0          1 1.19
## 58     0          0          0          1 1.35
## 59     0          0          0          1 1.54
## 60     0          0          0          1 1.48
## 61     0          0          0          0 1.79
## 62     0          0          0          0 2.05
## 63     0          0          0          0 1.85
## 64     0          0          0          0 1.95
```

Now that the data is properly formatted, we can begin our analysis.

Analysis

```
model <- lm(y ~ Rotar*Leg.Length*Leg.Width*Paperclip, data = df)
summary(model)
```

```
##
## Call:
## lm(formula = y ~ Rotar * Leg.Length * Leg.Width * Paperclip,
##     data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.25750 -0.10312 -0.00875  0.09188  0.47250
##
## Coefficients:
##                                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)                           1.91000    0.08578  22.266  < 2e-16 ***
## Rotar                                 0.02250    0.12131   0.185  0.85364
## Leg.Length                            0.01000    0.12131   0.082  0.93465
## Leg.Width                             0.04500    0.12131   0.371  0.71231
## Paperclip                            -0.52000    0.12131  -4.286  8.7e-05 ***
## Rotar:Leg.Length                     -0.00500    0.17156  -0.029  0.97687
## Rotar:Leg.Width                       0.17000    0.17156   0.991  0.32670
## Leg.Length:Leg.Width                 -0.02750    0.17156  -0.160  0.87332
## Rotar:Paperclip                      -0.00500    0.17156  -0.029  0.97687
## Leg.Length:Paperclip                  0.48000    0.17156   2.798  0.00738 **
## Leg.Width:Paperclip                   0.34750    0.17156   2.026  0.04839 *
## Rotar:Leg.Length:Leg.Width           -0.23250    0.24262  -0.958  0.34272
## Rotar:Leg.Length:Paperclip           -0.29250    0.24262  -1.206  0.23389
## Rotar:Leg.Width:Paperclip             0.48250    0.24262   1.989  0.05245 .
## Leg.Length:Leg.Width:Paperclip       -0.39500    0.24262  -1.628  0.11006
## Rotar:Leg.Length:Leg.Width:Paperclip -0.14750    0.34312  -0.430  0.66921
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1716 on 48 degrees of freedom
## Multiple R-squared:  0.7324, Adjusted R-squared:  0.6488
## F-statistic:  8.76 on 15 and 48 DF,  p-value: 3.345e-09
```

We would like to provide some interpretation of our model. However, before we do so, we see from our model output that none of the three and four way interaction effects appear to be statistically significant. Although, the three way interaction for the rotor, leg width, and paperclip has a p-value that is near 0.05. Therefore, in order to ease interpretation of the model, we would like to remove them from our modelling but must first gather more evidence that they are not statistically significant. In order to do this, we will conduct a nested F-test on the full factorial model and a model containing only two way interactions to ensure that the three and four way interactions are not statistically significant.

```r
model2 <- lm(y ~ Rotar*Leg.Length + Leg.Width*Paperclip +
               Rotar*Leg.Width + Rotar*Paperclip +
               Leg.Width*Leg.Length + Paperclip*Leg.Length, data = df)
anova(model2, model)
```

```
## Analysis of Variance Table
##
## Model 1: y ~ Rotar * Leg.Length + Leg.Width * Paperclip + Rotar * Leg.Width +
##     Rotar * Paperclip + Leg.Width * Leg.Length + Paperclip *
##     Leg.Length
## Model 2: y ~ Rotar * Leg.Length * Leg.Width * Paperclip
##   Res.Df    RSS Df Sum of Sq      F   Pr(>F)
## 1     53 2.0330
## 2     48 1.4128  5   0.62017 4.2141 0.002957 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Unfortunately, a nested F-test provided us with evidence that there is a statistically significant difference between the models estimating three and four way effects and the model that is not estimating them. Therefore, let us attempt the same nested F-test replacing the model of two way interactions with a model for three way interactions.

```r
model3 <- lm(y ~ Rotar*Leg.Length + Leg.Width*Paperclip +
               Rotar*Leg.Width + Rotar*Paperclip +
               Leg.Width*Leg.Length + Paperclip*Leg.Length +
               Rotar*Leg.Width*Paperclip + Leg.Length*Leg.Width*Paperclip +
               Rotar*Leg.Length*Paperclip +
               Rotar*Leg.Length*Leg.Width, data = df)
anova(model3, model)
```

```
## Analysis of Variance Table
##
## Model 1: y ~ Rotar * Leg.Length + Leg.Width * Paperclip + Rotar * Leg.Width +
##     Rotar * Paperclip + Leg.Width * Leg.Length + Paperclip *
##     Leg.Length + Rotar * Leg.Width * Paperclip + Leg.Length *
##     Leg.Width * Paperclip + Rotar * Leg.Length * Paperclip +
##     Rotar * Leg.Length * Leg.Width
## Model 2: y ~ Rotar * Leg.Length * Leg.Width * Paperclip
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1     49 1.4182
## 2     48 1.4128  1 0.0054391 0.1848 0.6692
```

Here, we see from the p-value that there is not statistically significant evidence that a four way interaction effect exists between our factors and that including the four way interaction in our model does not improve our predictive power. Therefore, we can continue our analysis by removing the four way interaction effect from our full factorial model in order to ease the interpretation of our model. Therefore, let us continue our interpretation of the model with the model containing all three way interaction effects. Let us print our model output.

```
summary(model3)
```

```
##
## Call:
## lm(formula = y ~ Rotar * Leg.Length + Leg.Width * Paperclip +
##     Rotar * Leg.Width + Rotar * Paperclip + Leg.Width * Leg.Length +
##     Paperclip * Leg.Length + Rotar * Leg.Width * Paperclip +
##     Leg.Length * Leg.Width * Paperclip + Rotar * Leg.Length *
##     Paperclip + Rotar * Leg.Length * Leg.Width, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.26672 -0.09391 -0.01250  0.09102  0.46328
##
## Coefficients:
##                                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)                     1.919219   0.082362  23.302  < 2e-16 ***
## Rotar                           0.004062   0.112528   0.036 0.971348
## Leg.Length                     -0.008438   0.112528  -0.075 0.940535
## Leg.Width                       0.026562   0.112528   0.236 0.814376
## Paperclip                      -0.538438   0.112528  -4.785 1.61e-05 ***
## Rotar:Leg.Length                0.031875   0.147334   0.216 0.829617
## Leg.Width:Paperclip             0.384375   0.147334   2.609 0.012012 *
## Rotar:Leg.Width                 0.206875   0.147334   1.404 0.166590
## Rotar:Paperclip                 0.031875   0.147334   0.216 0.829617
## Leg.Length:Leg.Width            0.009375   0.147334   0.064 0.949523
## Leg.Length:Paperclip            0.516875   0.147334   3.508 0.000977 ***
## Rotar:Leg.Width:Paperclip       0.408750   0.170127   2.403 0.020111 *
## Leg.Length:Leg.Width:Paperclip -0.468750   0.170127  -2.755 0.008209 **
## Rotar:Leg.Length:Paperclip     -0.366250   0.170127  -2.153 0.036284 *
## Rotar:Leg.Length:Leg.Width     -0.306250   0.170127  -1.800 0.077998 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1701 on 49 degrees of freedom
## Multiple R-squared:  0.7314, Adjusted R-squared:  0.6547
## F-statistic: 9.531 on 14 and 49 DF,  p-value: 1.14e-09
```

For all of our model analysis, we will use a significance threshold of 0.1.

We see from our model that a paper airplane with low settings for all four factors, our baseline variant, is expected to fly for 1.919219 seconds. If we change the rotor factor to high, we expect the paper airplane to fly 0.004062 seconds longer than the baseline model on average. However, this interaction is not statistically significant at all. If we maintain the baseline design and change the leg length factor to high, we expect the paper airplane to fly 0.008438 seconds shorter than the baseline model on average. Again however, this interaction is not statistically significant at all. If we maintain the baseline design and change the leg width factor to high, we expect the paper airplane to fly 0.026562 seconds longer than the baseline model on average. Again however, this interaction is not statistically significant at all. The only variant with a single high factor to result in a statistically significant change in flight time compared to the baseline model is when we maintain the baseline design but set the paperclip factor to high. We expect this variant to fly for 0.538438 seconds shorter than the baseline design.

If we keep the paperclip factor on high and additionally change the leg length factor to high while keeping the other two factors on low, we expect to see another statistically significant difference in the flight time of the helicopter that we expect to be 0.516875 seconds longer than if the paperclip was the only factor set to high. Additionally, we expect the flight time to be on average 0.516875 seconds longer than if the leg length was the only factor set to high. Instead, if we keep the paperclip factor on high and change the leg width factor to high while keeping the other two factors on low, we expect to see another statistically significant difference in the flight time of the helicopter that we expect to be 0.384375 seconds longer than if the paperclip was the only factor set to high. Additionally, we expect the flight time to be on average 0.384375 seconds longer than if the leg length was the only factor set to high. These are the only two way interactions that appear to be statistically significant.

Meanwhile, we will also provide an interpretation of the two way interaction effects that do not appear to be significant. Namely, the interaction effects between the rotor and leg length, between the rotor and leg width, between the leg length and leg width, and finally between the rotor and paperclip. While not statistically significant, our model predicts that on average if the rotor and leg length are high while the other two factors are low, the flight time of the helicopter will 0.031875 seconds longer than if only the rotor or leg length were high. Additionally, while not statistically significant, our model predicts that on average if the rotor and leg width are high while the other two factors are low, the flight time of the helicopter will 0.206875 seconds longer than if only the rotor or leg width were high. The interaction effect between the leg length and leg width is also not statistically significant according to our model. Despite not being statistically significant, our model predicts that the flight time of the helicopter will be 0.009375 seconds longer than if only the leg length or leg width were high while the other factors were low. Finally, while not statistically significant, our model predicts that on average if the rotor and paperclip are high while the other two factors are low, the flight time of the helicopter will 0.031875 seconds longer than if only the rotor or paperclip were high.

For our three way interactions, it seems as if all of the three way interactions appear to be significant, indicating that there is a significant change in the flight time of the helicopter when at least three of the factors are set to high in a variant. The interaction between rotor, leg width and paperclip has a coefficient of 0.408750, indicating that if all three of these factors are high, we expect the flight time of the helicopter to be 0.408750 seconds longer than if only two of these factors were high while the third one was low. Next, we see that the interaction between rotor, leg length and paperclip has a coefficient of -0.366250, indicating that if all three of these factors are high, we expect the flight time of the helicopter to be 0.366250 seconds shorter than if only two of these factors were high while the third one was low. We then see that the interaction between leg width, leg length and paperclip has a coefficient of -0.468750, indicating that if all three of these factors are high, we expect the flight time of the helicopter to be 0.468750 seconds shorter than if only two of these factors were high while the third one was low. Finally, we see that the interaction between leg width, leg length and paperclip has a coefficient of -0.306250, indicating that if all three of these factors are high, we expect the flight time of the helicopter to be 0.306250 seconds shorter than if only two of these factors were high while the third one was low.

Limitations

We also have a few limitations that we would like to discuss in our design. The first limitation that we have is that the helicopters were made by different individuals. The fact that two individuals were making different variants of helicopters could introduce some latent variables into the data that our model did not capture. For example, these could include the type of paper, the draft in the house, and other area specific factors. Additionally, there could be some bias introduced when measuring the flight time of the planes. One individual testing could have slower reaction times and therefore be slower to stop the stopwatch when the helicopter touches the ground. With such small numbers, a lack of precision in timing could have profound impacts on the model. Therefore, we would like to make note of these limitations in our design and methodology.

Conclusion

The main objectives of this analysis were to answer the three questions listed at the beginning of this paper.

1. The factor that seems to be the most important for making helicopters that fly longer appears to be the paperclip. When the paperclip is set to low, meaning that there is no paperclip on the helicopter, the helicopter tends to fly longer. However, when included in other interaction effects involving the width and/or length of the helicopter legs, the paperclip seems to have significant results on the flying time of the helicopter as well. There does not appear to be a distinct pattern for how the paperclip interacts with these other factors. However, it does appear that whenever there is an interaction between the paperclip and either the leg width or the leg length, there is a significant effect on flight time. These effects can be seen above.

2. While it does not appear that the effect of rotor length differs by leg width, the interaction between the rotor length, leg width, and paperclip appears to be significant. The interaction between the leg length, leg width, and the rotor also appears to be significant. These interactions do not give any evidence that the effect of rotor length varies due solely to the leg width, however. Instead, it implies that the effects of the leg length and leg width together vary the effect of the rotor. Additionally, it implies that effects of the rotor length vary dependent on the paperclip and leg width factors togehter. There is not a significant effect indicating that the effect of the rotor length varies due solely to the leg width.

3. From our model, it appears as if the variant where the rotar, paperclip, and leg width are set to high, while the leg length is set to low results in the longest flying helicopter. Our model estimates that this design would on average have a flying of time of 2.44328 seconds, which is much higher than any of the other variants that we estimated.