STA 322 Project 2

Alayna Binder & Katherine Zhong

2025-12-05

1 Experimental Design

We conducted a full 2^4 factorial experiment that included all sixteen combinations of the four helicopter design factors A (rotor length), B (leg lenth), C (leg width), and D (paper clip on leg). A full factorial structure was appropriate because our goal was to estimate the main effects of each design feature and evaluate whether any interactions influenced flight time, including interaction terms that might be aliased in a fractional factorial design. We collected six replicate drops for each design to obtain precise estimates of the factor effects while keeping the data collection workload manageable.

To reduce the possibility of ordering effects, we randomized the sequence of the sixteen design combinations before beginning data collection. After establishing this order, we also randomized which thrower completed each drop (among the six drops for each factor combination). Both throwers aimed to release the helicopters in a consistent manner, and random assignment ensures that any natural differences in release technique were distributed evenly across all treatments.

We treated thrower as a blocking variable. Although using a single thrower would theoretically eliminate person-to-person variation, it would also require one individual to perform all 96 drops. That workload increases the likelihood of fatigue or gradual shifts in release technique, which can introduce time-related variability that is difficult to control. We judged this within-person drift to be a greater risk than the small, consistent differences that might exist between two people. Because both project members were involved in data collection, blocking was the appropriate design choice. By structuring the experiment so that each thrower completed drops for every design combination, we created a fully crossed block that allowed us to account for person-specific differences directly and improve the precision of the estimated factor effects.

All drops were completed within a single session, with both experimenter present. We tested the helicopter drops one by one. While some within-session variation is possible, the number of drops per person was modest after splitting the workload, and the randomized allocation of treatments ensures that any minor temporal changes did not align with specific factor levels. We did not observe systematic shifts in flight time over the course of the session, so we treated within-session variation as negligible.

2 Analysis

2.1 Question 1

To identify which design factors influence flight time, we first examined the distributions of flight times across the two levels of each factor. The main effect boxplots below show a clear upward shift for factor A, with level 1 producing noticeably longer flights than level 0. Factor D shows the opposite pattern, with shorter flights at level 1. In contrast, factors B and C display only small differences between levels, and the variation within each level is larger than any apparent shift between them.

Main Effect Plots for Factors A–D
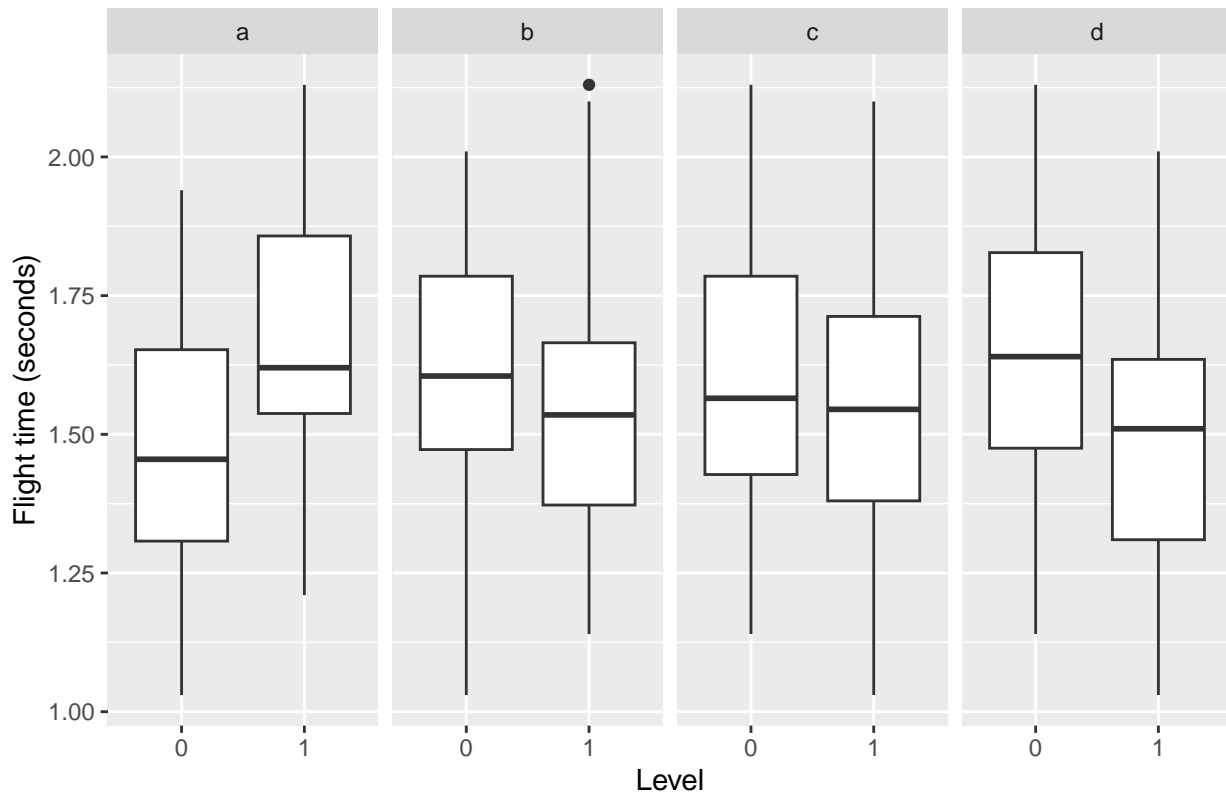
We then formalized these patterns using a linear model that included all four design factors while blocking on thrower. (We included `maker` as a variable in the data file but we did not use it in our model because both experimenters followed the instructions faithfully, and thus we don't expect the person who made the helicopter to make a difference in the flight time). A check of the model assumptions shows that the diagnostics are acceptable: the residuals vs. fitted plot shows no systematic curvature, supporting the linearity assumption; the Q-Q plot follows the reference line closely aside from minor tail deviations, indicating approximate normality; and the scale-location plot suggests that the spread of residuals remains fairly constant across fitted values, indicating no strong heteroscedasticity. The constant leverage plot also shows no observations with both large residuals and high leverage, so there are no influential points that would compromise the model.
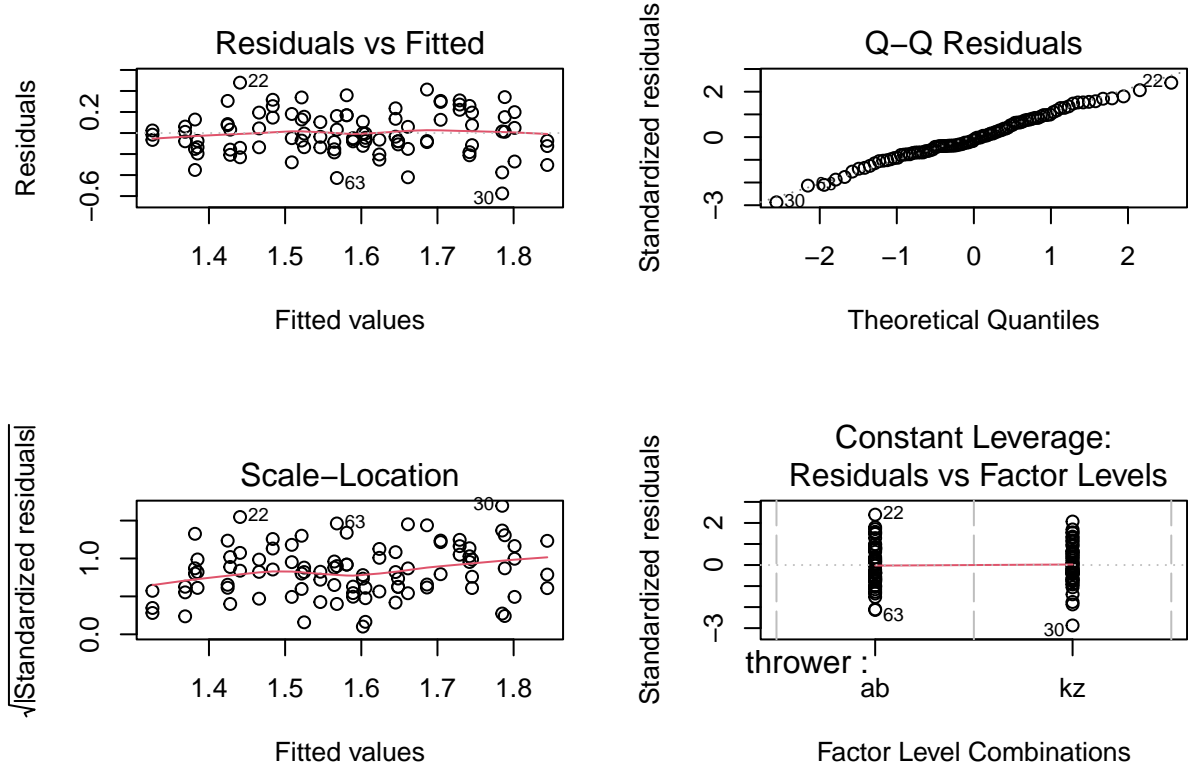
The estimated effects and confidence intervals are shown in the table below. Factor A had a positive and statistically significant effect, increasing the mean flight time by about 0.22 seconds. Factor D had a negative and significant effect, decreasing the mean flight time by about 0.14 seconds. The confidence intervals for both effects were entirely away from zero. Factors B and C had small estimated effects with confidence intervals that included zero, indicating little evidence that they meaningfully affect flight time.

Table 1: Main Effects Model Results

term	estimate	conf.low	conf.high	std.error	statistic	p.value
(Intercept)	1.624	1.521	1.726	0.052	31.411	0.000
a	0.220	0.137	0.304	0.042	5.222	0.000
b	-0.056	-0.140	0.028	0.042	-1.323	0.189
c	-0.043	-0.127	0.041	0.042	-1.017	0.312
d	-0.140	-0.224	-0.056	0.042	-3.317	0.001
throwerkz	-0.059	-0.143	0.025	0.042	-1.402	0.164

Both the exploratory plots and the fitted model point to the same conclusion. Factors A and D are the most influential design components. Factor A corresponds to rotor length, and the model indicates that longer rotors tend to increase flight time. Factor D corresponds to whether a paper clip is attached, and the negative effect estimate suggests that adding a paper clip shortens the flight. In contrast, the variables associated with leg length (factor B) and leg width (factor C) show minimal impact on flight duration under this experimental setup.

Although the exploratory plots and main-effects model indicate that only factors A and D meaningfully influence flight time, these conclusions rely on a model that does not include interaction terms. Because factors in a full factorial experiment may contribute through interactions even when their main effects appear small, we also compared the full interaction model to reduced models that removed each factor and all of its associated interaction terms (see Appendix for code). These nested ANOVA comparisons quantify the total contribution of each factor to model fit. Removing factor A or factor D led to substantial increases in

residual error (Δadj-$R^2 = 0.259$ and 0.312, respectively), reaffirming that both factors are highly influential. Removing factors B or C also significantly worsened the model (Δadj-$R^2 = 0.193$ and 0.150), suggesting that they participate in interaction patterns even though their main effects alone were not statistically significant. All p-values for each of the nested F-tests between the full model and the models with factors involving a, b, c, or d removed were very small (all of them smaller than .01). Thus, while the simple main-effects model highlights A and D as the most important design features, the broader model-comparison results show that all four factors contribute to explaining flight time to varying degrees, with D and A having the strongest overall influence. Full details and output tables for these comparisons are provided in the Appendix.

2.2 Question 2

We use a linear regression model to investigate this potential interaction effect between rotor length (A) and leg width (C).

First, we check the model assumptions of a full model with all main effects, all two-way, three-way and four-way interactions between A, B, C, D, and `thrower` as the block variable. We do so by checking the residual vs. fitted plot, Q-Q plot, scale-location plot, and residuals vs. leverage plot.

We choose to have the block variable as a fixed effect in the model since we do not expect the thrower effect to be different based on the different helicopter features.
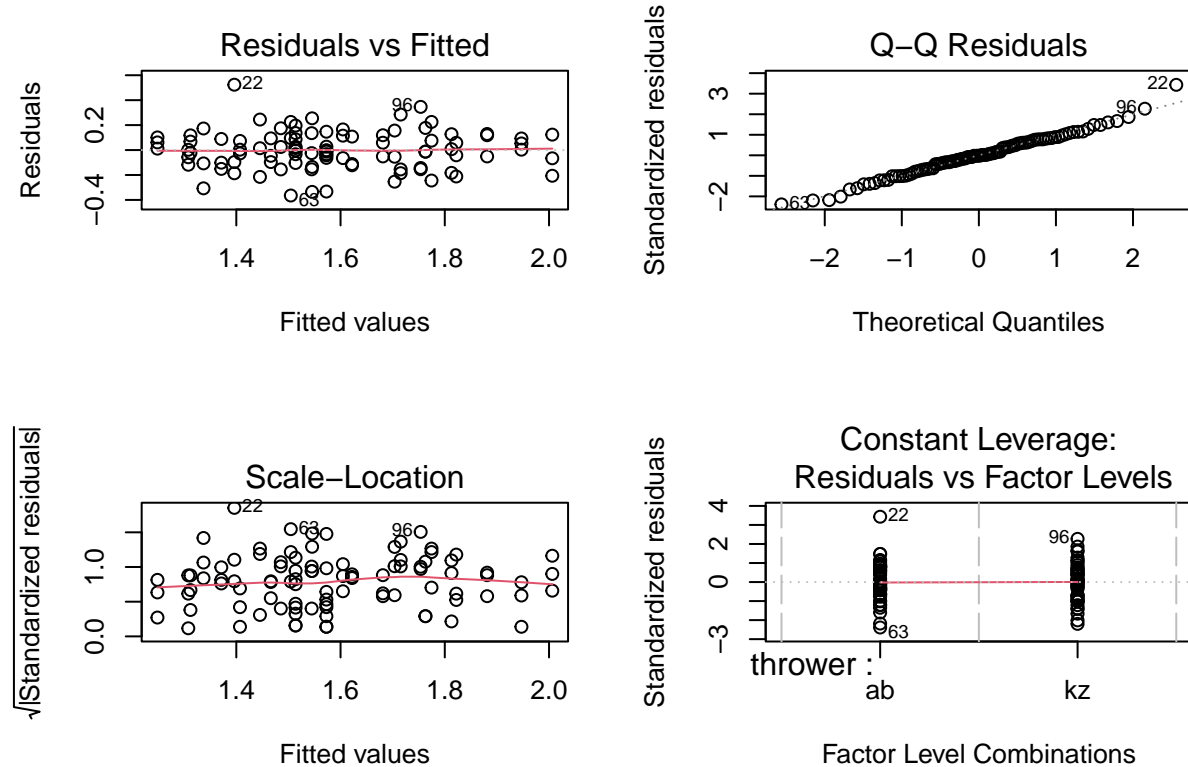
Overall, the model diagnostics suggest that the linear regression model is appropriate for these data. The residuals appear to be approximately normally distributed, and the residuals vs. fitted plot shows no clear patterns that would indicate non-linearity or heteroscedasticity. The Q-Q plot further supports the normality assumption since the residuals fall close to the diagonal reference line, meaning their distribution aligns well with what we would expect under normality, with no strong skewness or unusual tail behavior. The scale-location plot also shows a fairly horizontal trend, which indicates that the variance of the residuals remains roughly constant across levels of the fitted values. While a few potential outliers are present, the constant leverage plot suggests that they do not appear to have a substantial influence on the model or meaningfully compromise the overall fit.

Since we are interested in the two-way interaction between rotor length (A) and leg width (C), we fit a reduced model that includes only the main effects and all two-way interaction terms, along with the blocking variable `thrower`.

Then, we compare whether the model with only main effects and two-way interactions is significantly different from the full model with all interaction terms using an ANOVA test.

Table 2: ANOVA Comparison of Reduced and Full Models

Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
84	2.912025	NA	NA	NA	NA
79	2.236583	5	0.6754417	4.771554	0.0007321

The ANOVA test comparing the full model with all interaction terms to the reduced model with only main effects and two-way interactions yields a p-value of 0.0007. The p-value is less than 0.05, indicating that there is a significant difference between the two models, suggesting that the additional higher-order interaction terms in the full model does provide a better fit to the data.

Additionally, we fit a model with all main effects, two-way interaction effects, three-way interaction effects, and the blocking variable `thrower` to see if including the full model significantly improves the model fit compared to the three-way interaction model.

Table 3: ANOVA Comparison of Three-Way Interaction and Full Models

Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
80	2.250600	NA	NA	NA	NA
79	2.236583	1	0.0140167	0.495093	0.4837329

The ANOVA test comparing the three-way interaction model to the full model yields a p-value of 0.4837 Since the p-value is greater than 0.05, this indicates that including the four-way interaction effect does not significantly improve the model fit compared to the three-way interaction model. Therefore, we will use the three-way interaction model for further analysis.

Lastly, for confirmation, we can also compare the two-way interaction model to the three-way interaction model using an ANOVA test.

Table 4: ANOVA Comparison of Two-Way and Three-Way Interaction Models

Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
84	2.912025	NA	NA	NA	NA
80	2.250600	4	0.661425	5.877766	0.0003371

Since p-value = 0.0003 > 0.05, this indicates that including the three-way interactions model does significantly improve the model fit compared to the two-way interaction model.

Table 5: Linear Model with Three-Way Interaction Results

term	estimate	conf.low	conf.high	std.error	statistic	p.value
(Intercept)	1.454	1.318	1.590	0.068	21.237	0.000
a	0.163	-0.018	0.343	0.091	1.794	0.077
b	0.063	-0.118	0.243	0.091	0.690	0.492
c	0.321	0.141	0.501	0.091	3.542	0.001
d	0.239	0.059	0.419	0.091	2.640	0.010
throwerkz	-0.059	-0.127	0.009	0.034	-1.728	0.088
a:b	0.315	0.079	0.551	0.119	2.656	0.010
a:c	-0.175	-0.411	0.061	0.119	-1.476	0.144
a:d	0.013	-0.223	0.249	0.119	0.112	0.911
b:c	-0.305	-0.541	-0.069	0.119	-2.572	0.012
b:d	-0.397	-0.633	-0.161	0.119	-3.345	0.001
c:d	-0.630	-0.866	-0.394	0.119	-5.312	0.000
a:b:c	-0.010	-0.283	0.263	0.137	-0.073	0.942
a:b:d	-0.265	-0.538	0.008	0.137	-1.935	0.057
a:c:d	0.200	-0.073	0.473	0.137	1.460	0.148
b:c:d	0.575	0.302	0.848	0.137	4.199	0.000

Based on our regression model results above, the estimated interaction effect of AC is -0.175 with a p-value of 0.144. The estimated effect size is small and $p > 0.05$, indicating insufficient evidence to support the claim of a significant interaction effect between rotor length (A) and leg width (C) on flight time. The confidence interval for the AC interaction term also includes zero (-0.411, 0.061), further confirming that the interaction effect is not statistically significant.

2.3   Question 3

To determine the helicopter design that maximizes flight time, we estimated the average flight time for each of the sixteen combinations of factors A through D using the three-way interaction model selected in Question 2, which provided the best overall fit.

We obtained these estimates using the `emmeans` package. `emmeans` computes the model-based mean for each factor combination by evaluating the fitted regression model at the specified factor levels and averaging over the blocking variable, while properly incorporating all main effects and interaction terms. We chose this method because it provides consistent and reliable predictions without repeatedly altering reference levels in the data. A comparison with the reference-level approach, described in the appendix, confirms that both methods produce the same results.

Table 6: Estimated average flight time (seconds) for all 16 factor combinations

Levels (A,B,C,D)	Estimate (s)	95% CI low	95% CI high
(1, 1, 0, 0)	1.965	1.833	2.097
(1, 0, 0, 1)	1.840	1.708	1.972
(1, 1, 1, 0)	1.795	1.663	1.927
(0, 0, 1, 0)	1.745	1.613	1.877
(1, 0, 1, 0)	1.733	1.601	1.865
(0, 0, 0, 1)	1.664	1.532	1.796
(1, 0, 0, 0)	1.587	1.455	1.719
(1, 1, 0, 1)	1.555	1.423	1.687
(1, 0, 1, 1)	1.555	1.423	1.687
(1, 1, 1, 1)	1.531	1.399	1.663
(0, 1, 1, 0)	1.503	1.371	1.635
(0, 1, 0, 0)	1.487	1.355	1.619
(0, 0, 0, 0)	1.425	1.293	1.557
(0, 0, 1, 1)	1.355	1.223	1.487
(0, 1, 0, 1)	1.330	1.198	1.462
(0, 1, 1, 1)	1.290	1.158	1.422

The longest predicted flight time was obtained when A = 1, B = 1, C = 0, and D = 0, corresponding to long rotors, long legs, narrow leg width, and no paper clip. This design had an estimated mean flight time of 1.965 seconds (95% CI: 1.833, 2.097). Therefore, this combination is our recommended configuration for maximizing helicopter flight duration.

3 Appendix

3.1 Data dictionary

Here are the variables in our data file "STA322-project2-helicopter - Sheet1.csv":

(Each row corresponds to one drop of a helicopter)

- combination – the factor combination of the helicopter, e.g. 1, a, b, . . . , abcd
- a – Rotor length (0 - low - 7.5cm, 1 - high - 8.5cm)
- b – Leg length (0 - low - 7.5cm, 1 - high - 12cm)
- c – Leg width (0 - low - 3.2cm, 1 - high - 5.0cm)
- d – Paper clip on leg (0 - low - No, 1 - high - Yes)
- thrower – person who performed the drop (kz – Katherine, ab - Alayna)
- time – flight time (in seconds) for each drop
- order – the randomized sequence/order each combination of helicopter was tested, e.g. helicopter with factor combination "d" was tested first out of 16 combinations, so order = 1 for "d"
- maker – person who made the helicopter (kz – Katherine, ab - Alayna)

3.2 Additional explanation about Q1 and Q3

To supplement our Q1 analysis, we evaluated the importance of each factor by comparing the full interaction model to reduced models that omitted one factor and all interaction terms containing that factor. This approach assesses the total contribution of each factor (main effect + interactions), addressing the possibility that some factors may matter primarily through interactions even if their main effects appear small. Table A1 presents the nested ANOVA (partial F) results for these model comparisons, and Table A2 reports the corresponding changes in R² and adjusted R².

Table 7: Table A1. Nested ANOVA summary for removing each factor

Factor removed	Df (added)	Sum of Sq (added)	F	Pr(>F)
A	7	1.561	7.927	0.0e+00
B	7	1.219	6.189	8.0e-06
C	7	0.995	5.050	8.9e-05
D	7	1.837	9.329	0.0e+00

Table 8: Table A2. Change in R² and adjusted R² when removing each factor

Factor removed	Change in R²	Change in Adj R²
A	0.274	0.262
B	0.214	0.196
C	0.175	0.153
D	0.323	0.315

These results provide the detailed evidence referred to in Q1. The nested ANOVA tests (Table A1) show how much model fit deteriorates when a factor and all its associated interactions are removed, and the R^2 summaries (Table A2) quantify the corresponding loss in explained variability. Using our knowledge from Q2 that the three-way interaction model was the best fit, we also ran this analysis using that model. The

p-values and R^2/adjusted-R^2 values were slightly different (differences negligible), but the ultimate conclusion was the same (i.e. D and A most influential, B and C main effects alone not influential).

For Question 3, we also used a manual approach that involved changing factor reference levels and extracting the model intercept to obtain the predicted means. This method produced the same results as `emmeans`, which confirms that both procedures are consistent with the fitted model. We included the manual approach to verify our predictions, but `emmeans` is more practical in application because it computes all combinations directly and avoids the repeated refitting required by the manual method.

Average flight time for factor combination (1) with factors (-, -, -, -):

Estimate: 1.454 seconds

95% CI: (1.318 , 1.59) seconds

Average flight time for factor combination a with factors (+, -, -, -):
Estimate: 1.617 seconds
95% CI: (1.48 , 1.753) seconds

Average flight time for factor combination b with factors (-, +, -, -):
Estimate: 1.517 seconds
95% CI: (1.38 , 1.653) seconds

Average flight time for factor combination c with factors (-, -, +, -):
Estimate: 1.775 seconds
95% CI: (1.639 , 1.911) seconds

Average flight time for factor combination d with factors (-, -, -, +):
Estimate: 1.693 seconds
95% CI: (1.557 , 1.83) seconds

Average flight time for factor combination ab with factors (+, +, -, -):
Estimate: 1.994 seconds
95% CI: (1.858 , 2.13) seconds

Average flight time for factor combination ac with factors (+, -, +, -):
Estimate: 1.762 seconds
95% CI: (1.626 , 1.899) seconds

Average flight time for factor combination ad with factors (+, -, -, +):
Estimate: 1.869 seconds
95% CI: (1.733 , 2.005) seconds

Average flight time for factor combination bc with factors (-, +, +, -):
Estimate: 1.533 seconds
95% CI: (1.396 , 1.669) seconds

Average flight time for factor combination bd with factors (-, +, -, +):
Estimate: 1.359 seconds
95% CI: (1.223 , 1.495) seconds

Average flight time for factor combination cd with factors (-, -, +, +):
Estimate: 1.384 seconds
95% CI: (1.248 , 1.52) seconds

Average flight time for factor combination abc with factors (+, +, +, -):
Estimate: 1.825 seconds
95% CI: (1.689 , 1.961) seconds

Average flight time for factor combination abd with factors (+, +, -, +):
Estimate: 1.585 seconds
95% CI: (1.449 , 1.721) seconds

```
## Average flight time for factor combination acd with factors (+, -, +, +):
## Estimate: 1.585 seconds
## 95% CI: ( 1.449 ,   1.721 ) seconds

## Average flight time for factor combination bcd with factors (-, +, +, +):
## Estimate: 1.32 seconds
## 95% CI: ( 1.184 ,   1.456 ) seconds

## Average flight time for factor combination abcd with factors (+, +, +, +):
## Estimate: 1.561 seconds
## 95% CI: ( 1.425 ,   1.697 ) seconds
```

3.3 Code

Setup

```r
library(tidyverse)
library(broom)
library(knitr)
library(kableExtra)
library(emmeans)
library(dplyr)
library(tidyr)
```

Load data

```r
data <- read_csv("STA322-project2-helicopter - Sheet1.csv")
```

Data wrangling

```r
df <- data |>
  mutate(
    # treat a-d as 0/1 indicators for the four design factors
    across(a:d, as.integer),

    # factors for people
    thrower = factor(thrower),
    maker = factor(maker),

    # combination as a factor, ordered by how it appears
    combination = factor(combination, levels = unique(combination))
  )

# summary(df$time)
```

3.3.1 Q1

Main effects box plot

```r
main_effect_summaries <- df |>
  pivot_longer(a:d, names_to = "factor", values_to = "level") |>
  group_by(factor, level) |>
  summarise(
    mean_time = mean(time),
    sd_time = sd(time),
    n = n(),
    .groups = "drop"
  )
```

10

```
# main_effect_summaries

df |>
  pivot_longer(a:d, names_to = "factor", values_to = "level") |>
  mutate(level = factor(level)) |>
  ggplot(aes(x = level, y = time)) +
  geom_boxplot() +
  facet_wrap(~ factor, nrow = 1) +
  labs(
    title = "Main Effect Plots for Factors A-D",
    x = "Level",
    y = "Flight time (seconds)"
  )
```

Fit main effects model with blocked variable and check model assumptions

```
fit_q1 <- lm(time ~ a + b + c + d + thrower, data = df)

par(mfrow = c(2, 2))
plot(fit_q1)
```

Display main effects model results

```
# summary(fit_q1)

q1_table <- tidy(fit_q1, conf.int = TRUE) |>
  select(term, estimate, conf.low, conf.high, std.error, statistic, p.value)

q1_table |>
  kable(digits = 3, caption = "Main Effects Model Results") |>
  kable_classic(full_width = FALSE) |>
  kableExtra::kable_styling(latex_options = "HOLD_position")
```

Supplementary analysis

```
# Full model from Q2
full <- fit_q2_full
factors <- c("a","b","c","d")

res <- map_dfr(factors, function(f) {
  others <- setdiff(factors, f)
  fmla <- as.formula(paste0("time ~ (", paste(others, collapse = " + "), ")^3 + thrower"))
  red <- lm(fmla, data = df)
  a <- anova(red, full)
  a_df <- as.data.frame(a)
  diff_row <- a_df[2, ]

  r_full <- summary(full)$r.squared
  adjr_full <- summary(full)$adj.r.squared
  r_red <- summary(red)$r.squared
  adjr_red <- summary(red)$adj.r.squared

  tibble(
    Factor = toupper(f),
    Df_added = diff_row[[3]],
```

```
    SumSq_added = diff_row[[4]],
    F_stat = diff_row[[5]],
    p_value = diff_row[[6]],
    delta_R2 = r_full - r_red,
    delta_AdjR2 = adjr_full - adjr_red
  )
})

# Table A1: ANOVA summary
res %>%
  select(Factor, Df_added, SumSq_added, F_stat, p_value) %>%
  rename(
    `Factor removed` = Factor,
    `Df (added)` = Df_added,
    `Sum of Sq (added)` = SumSq_added,
    `F` = F_stat,
    `Pr(>F)` = p_value
  ) %>%
  kable(
    digits = c(0,0,3,3,6),
    caption = "Table A1. Nested ANOVA summary for removing each factor"
  ) %>%
  kable_styling(full_width = FALSE, latex_options = "HOLD_position")

cat("\n\n")

# Table A2: R-sq and adj-R-sq
res %>%
  select(Factor, delta_R2, delta_AdjR2) %>%
  rename(
    `Factor removed` = Factor,
    `Change in R²` = delta_R2,
    `Change in Adj R²` = delta_AdjR2
  ) %>%
  mutate(across(c(`Change in R²`, `Change in Adj R²`), ~ round(., 3))) %>%
  kable(
    digits = 3,
    caption = "Table A2. Change in R² and adjusted R² when removing each factor"
  ) %>%
  kable_styling(full_width = FALSE, latex_options = "HOLD_position")
```

3.3.2 Q2

Fit full model and check model assumptions

```
# Make linear model with all effects plus block variable
fit_q2_full <- lm(time ~ a * b * c * d + thrower, data = df)

# summary(fit_q2_full)
# Check model fit with full linear model
par(mfrow = c(2, 2))
plot(fit_q2_full)
```

Fit reduced model with main effects and two-way interactions

```
# Fit reduced model with main effects and two-way interactions
fit_q2_two_way <- lm(time ~ (a + b + c + d)^2 + thrower, data = df)

# summary(fit_q2_two_way)
```

ANOVA to compare full model with two-way interaction model

```
# ANOVA test to compare full model and reduced model
anova_res <- anova(fit_q2_two_way, fit_q2_full)

kable(anova_res, caption = "ANOVA Comparison of Reduced and Full Models") %>%
  kable_styling(full_width = FALSE, position = "center") |>
  kableExtra::kable_styling(latex_options = "HOLD_position")
```

Fit three-way interaction model; ANOVA to compare with full model

```
# Fit model with main effects, two-way and three-way interactions
fit_q2_three_way <- lm(time ~ (a + b + c + d)^3 + thrower, data = df)

# ANOVA test to compare three-way interaction model and full model
anova2 <- anova(fit_q2_three_way, fit_q2_full)

kable(anova2, caption = "ANOVA Comparison of Three-Way Interaction and Full
      Models") %>%
  kable_styling(full_width = FALSE, position = "center") |>
  kableExtra::kable_styling(latex_options = "HOLD_position")
```

ANOVA to compare two-way interaction model with three-way interaction model

```
# ANOVA test to compare three-way interaction model and two way model
anova3 <- anova(fit_q2_two_way, fit_q2_three_way)

kable(anova3, caption = "ANOVA Comparison of Two-Way and Three-Way
      Interaction Models") %>%
  kable_styling(full_width = FALSE, position = "center") |>
  kableExtra::kable_styling(latex_options = "HOLD_position")
```

Display final model (model with three-way interactions) output

```
q2_table <- tidy(fit_q2_three_way, conf.int = TRUE) |>
  select(term, estimate, conf.low, conf.high, std.error, statistic, p.value)

# Model output in tidy table
q2_table |>
  kable(digits = 3, caption = "Linear Model with Three-Way Interaction Results") |>
  kable_classic(full_width = FALSE) |>
  kableExtra::kable_styling(latex_options = "HOLD_position")
```

3.3.3   Q3

Get estimated average flight time for all 16 factor combinations using `emmeans`

```
# Get estimated marginal means for all 16 combinations
emm_all <- emmeans(fit_q2_three_way, ~ a * b * c * d, infer = TRUE, level = 0.95)

emm_table <- as.data.frame(emm_all) %>%
  rename(estimate = emmean, conf.low = lower.CL, conf.high = upper.CL) %>%
```

```r
  mutate(
    factors = paste0("(", a, ", ", b, ", ", c, ", ", d, ")")
  ) %>%
  select(factors, estimate, conf.low, conf.high) %>%
  arrange(desc(estimate))

kable(emm_table, digits = 3,
      col.names = c("Levels (A,B,C,D)", "Estimate (s)", "95% CI low", "95% CI high"),
      caption = "Estimated average flight time (seconds) for all
      16 factor combinations") |>
      kableExtra::kable_styling(latex_options = "HOLD_position")
```

Manually calculated the average flight time for factor combination (1), with factors (-, -, -, -):

```r
# Get average flight time for factor combination (1)
avg_flight_time_1 <- tidy(fit_q2_three_way, conf.int = TRUE) |>
  filter(term == "(Intercept)") |>
  select(estimate, conf.low, conf.high)

# print the average flight time and 95% CI for factor combination (1)

cat("Average flight time for factor combination (1) with factors (-, -, -, -):\n")
cat("Estimate:", round(avg_flight_time_1$estimate, 3), "seconds\n")
cat("95% CI: (", round(avg_flight_time_1$conf.low, 3), ", ",
    round(avg_flight_time_1$conf.high, 3), ") seconds \n")
```

Write function to calculate the average flight times for the 15 other combinations

```r
estimate_avg_flight_time <- function(df_modified, combination_name, factors_desc) {
  # Fit linear model with new factor levels.
  # (a + b + c + d)^3 fits all main effects, 2-way, and 3-way interactions.
  fit_q3 <- lm(time ~ (a + b + c + d)^3 + thrower, data = df_modified)

  # Get average flight time (intercept) with 95% CI
  avg_flight_time <- tidy(fit_q3, conf.int = TRUE) |>
    filter(term == "(Intercept)") |>
    select(estimate, conf.low, conf.high)

  # print the average flight time and 95% CI
  cat(sprintf("Average flight time for factor combination %s with factors %s:\n",
              combination_name, factors_desc))
  cat("Estimate:", round(avg_flight_time$estimate, 3), "seconds\n")
  cat("95% CI: (", round(avg_flight_time$conf.low, 3), ", ",
      round(avg_flight_time$conf.high, 3), ") seconds \n\n")
}

# --- 2) Average flight time for factor combination a, with factors (+, -, -, -): ---
df_a <- df |>
  mutate(
    a = ifelse(a == 1, 0, 1), b = ifelse(b == 0, 0, 1),
    c = ifelse(c == 0, 0, 1), d = ifelse(d == 0, 0, 1)
  )
estimate_avg_flight_time(df_a, "a", "(+, -, -, -)")

# --- 3) Average flight time for factor combination b, with factors (-, +, -, -): ---
```

```r
df_b <- df |>
  mutate(
    a = ifelse(a == 0, 0, 1), b = ifelse(b == 1, 0, 1),
    c = ifelse(c == 0, 0, 1), d = ifelse(d == 0, 0, 1)
  )
estimate_avg_flight_time(df_b, "b", "(-, +, -, -)")

# --- 4) Average flight time for factor combination c, with factors (-, -, +, -): ---
df_c <- df |>
  mutate(
    a = ifelse(a == 0, 0, 1), b = ifelse(b == 0, 0, 1),
    c = ifelse(c == 1, 0, 1), d = ifelse(d == 0, 0, 1)
  )
estimate_avg_flight_time(df_c, "c", "(-, -, +, -)")

# --- 5) Average flight time for factor combination d, with factors (-, -, -, +): ---
df_d <- df |>
  mutate(
    a = ifelse(a == 0, 0, 1), b = ifelse(b == 0, 0, 1),
    c = ifelse(c == 0, 0, 1), d = ifelse(d == 1, 0, 1)
  )
estimate_avg_flight_time(df_d, "d", "(-, -, -, +)")

# --- 6) Average flight time for factor combination ab, with factors (+, +, -, -): ---
df_ab <- df |>
  mutate(
    a = ifelse(a == 1, 0, 1), b = ifelse(b == 1, 0, 1),
    c = ifelse(c == 0, 0, 1), d = ifelse(d == 0, 0, 1)
  )
estimate_avg_flight_time(df_ab, "ab", "(+, +, -, -)")

# --- 7) Average flight time for factor combination ac, with factors (+, -, +, -): ---
df_ac <- df |>
  mutate(
    a = ifelse(a == 1, 0, 1), b = ifelse(b == 0, 0, 1),
    c = ifelse(c == 1, 0, 1), d = ifelse(d == 0, 0, 1)
  )
estimate_avg_flight_time(df_ac, "ac", "(+, -, +, -)")

# --- 8) Average flight time for factor combination ad, with factors (+, -, -, +): ---
df_ad <- df |>
  mutate(
    a = ifelse(a == 1, 0, 1), b = ifelse(b == 0, 0, 1),
    c = ifelse(c == 0, 0, 1), d = ifelse(d == 1, 0, 1)
  )
estimate_avg_flight_time(df_ad, "ad", "(+, -, -, +)")

# --- 9) Average flight time for factor combination bc, with factors (-, +, +, -): ---
df_bc <- df |>
  mutate(
    a = ifelse(a == 0, 0, 1), b = ifelse(b == 1, 0, 1),
    c = ifelse(c == 1, 0, 1), d = ifelse(d == 0, 0, 1)
  )
```

```r
estimate_avg_flight_time(df_bc, "bc", "(-, +, +, -)")

# --- 10) Average flight time for factor combination bd, with factors (-, +, -, +): ---
df_bd <- df |>
  mutate(
    a = ifelse(a == 0, 0, 1), b = ifelse(b == 1, 0, 1),
    c = ifelse(c == 0, 0, 1), d = ifelse(d == 1, 0, 1)
  )
estimate_avg_flight_time(df_bd, "bd", "(-, +, -, +)")


# --- 11) Average flight time for factor combination cd, with factors (-, -, +, +): ---
df_cd <- df |>
  mutate(
    a = ifelse(a == 0, 0, 1), b = ifelse(b == 0, 0, 1),
    c = ifelse(c == 1, 0, 1), d = ifelse(d == 1, 0, 1)
  )
estimate_avg_flight_time(df_cd, "cd", "(-, -, +, +)")


# --- 12) Average flight time for factor combination abc, with factors (+, +, +, -): ---
df_abc <- df |>
  mutate(
    a = ifelse(a == 1, 0, 1), b = ifelse(b == 1, 0, 1),
    c = ifelse(c == 1, 0, 1), d = ifelse(d == 0, 0, 1)
  )
estimate_avg_flight_time(df_abc, "abc", "(+, +, +, -)")


# --- 13) Average flight time for factor combination abd, with factors (+, +, -, +): ---
df_abd <- df |>
  mutate(
    a = ifelse(a == 1, 0, 1), b = ifelse(b == 1, 0, 1),
    c = ifelse(c == 0, 0, 1), d = ifelse(d == 1, 0, 1)
  )
estimate_avg_flight_time(df_abd, "abd", "(+, +, -, +)")


# --- 14) Average flight time for factor combination acd, with factors (+, -, +, +): ---
df_acd <- df |>
  mutate(
    a = ifelse(a == 1, 0, 1), b = ifelse(b == 0, 0, 1),
    c = ifelse(c == 1, 0, 1), d = ifelse(d == 1, 0, 1)
  )
estimate_avg_flight_time(df_acd, "acd", "(+, -, +, +)")


# --- 15) Average flight time for factor combination bcd, with factors (-, +, +, +): ---
df_bcd <- df |>
  mutate(
    a = ifelse(a == 0, 0, 1), b = ifelse(b == 1, 0, 1),
    c = ifelse(c == 1, 0, 1), d = ifelse(d == 1, 0, 1)
  )
```

```r
estimate_avg_flight_time(df_bcd, "bcd", "(-, +, +, +)")


# --- 16) Average flight time for factor combination abcd, with factors (+, +, +, +): ---
df_abcd <- df |>
  mutate(
    a = ifelse(a == 1, 0, 1), b = ifelse(b == 1, 0, 1),
    c = ifelse(c == 1, 0, 1), d = ifelse(d == 1, 0, 1)
  )
estimate_avg_flight_time(df_abcd, "abcd", "(+, +, +, +)")
```