Paper Airplane Factorial Design

Lucas Childs

2025-03-14

Introduction

Paper airplanes have long been a fun and creative craft to engage children, but their flight dynamics also tie into real principles of aerodynamics. The factors affecting the flight of paper airplanes like weight distribution, mass of the plane, and drag, parallel those affecting real aircraft. Understanding how these changes made to the paper airplane's design impact its flight can offer insights into how small alterations in mass, aerodynamics, and fuel amount influence the flight distance of real airplanes [1, 2].

This study aims to analyze how the placement of paper clips on a paper airplane-at the nose, middle, or rear-affect its flight distance. The weight distribution, altered with paper clip placement, can significantly affect the plane's center of gravity, lift, and drag, causing resistance on the aircraft as well as risk for backflip when the aircraft is not weighted sufficiently in the nose [3, 4]. By testing different architectures and with statistical analysis, this experiment will provide insights into the optimal weight distribution to maximize flight distance.

Methods

The design of this experiment was a full factorial design with 3 factors, each with two levels: the attachment of a paper clip on the nose, middle, and rear of the plane, with levels 0 and 1, where 1 represents the presence of the paper clip corresponding to its factor. To collect data, the paper airplane was thrown along a measuring tape with different treatments according to all 8 combinations of the factors: no paper clips, nose paper clip, middle paper clip, rear paper clip, nose and middle paper clip, nose and rear paper clip, middle and rear paper clip, and nose and middle and rear paper clip. Each distance that the airplane flew under the specific conditions was recorded into an Excel spreadsheet in inches. The flight distance was determined by the spot the airplane first hit the ground, recorded to my best ability, since my vantage point was usually 14 feet away from this spot. The order of data collection was randomized with the base R `sample` function. Once I listed out 12 full replicates of the 8 possible factor combinations (determined through sample size calculations & recommended by Professor Chi) into Excel, the random sample shuffled the order of these factors before the data collection started. Then I wrote the shuffled data into an Excel file, added the `Distance` column, and loaded this data back into R.

Photos of the data collection setup and airplane:

```r
library(knitr)

include_graphics("data_setup.jpeg")
```

```
include_graphics("airplane.jpeg")
```

The question to be answered was whether or not any combination of placing a paper clip on the nose, middle, and rear of a paper airplane has any effect on flight distance. For statistical analysis, a linear regression model was used as well as a permutation test for a non-parametric method.

For a linear regression model, the data is assumed to be normally distributed, not structured, and have equal or constant variance (homoscedasticity). In this respect, it is possible that the data did not follow a normal distribution and homoscedasticity was not maintained due to inconsistencies in my throwing technique, physical fatigue, and deterioration of the paper airplane. A thorough analysis of whether or not these assumptions were met will be given in the Results section.

A technical issue that arose during the experiment was that after repeated trials, the nose of the paper airplane became crumpled, since it often hit the ground nose first. Additionally, a few flights ended inconclusively due to the interference of a counter-top next to the flight zone in which the trial for that factor or combination of factors was repeated. Moreover, I initially was not consistent with my tosses, potentially skewing data in the beginning of the experiment, however my throwing technique improved after several replicates.

Results

Sample size calculation via pilot study:

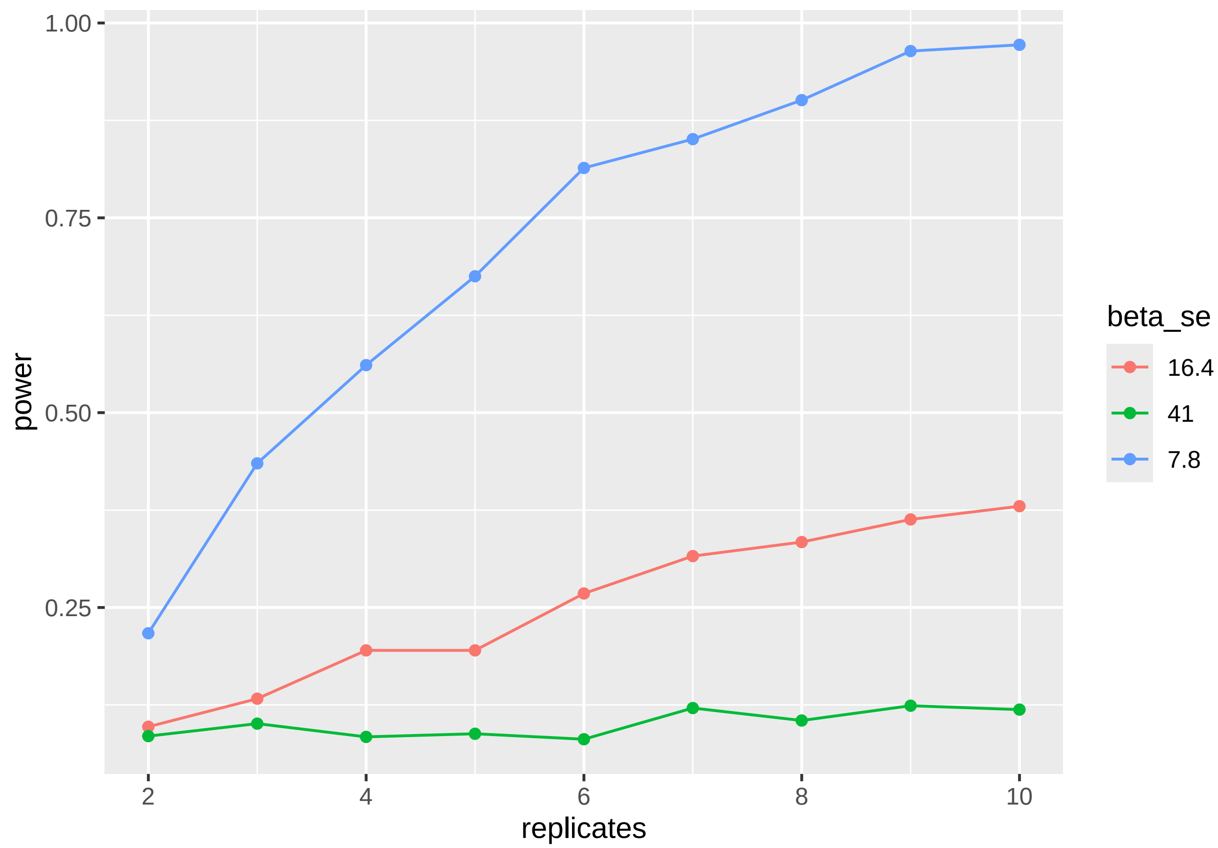
```r
set.seed(2025745)
library(ggplot2)
```

```r
library(readxl)
library(writexl)
source("power_factorial_23.R")

# Pilot study with 2 replicates
# Randomization of the order of data collection:

# tmp_pilot <- read_excel("tmp_pilot")
# shuffled_tmp_pilot <- tmp_pilot[sample(nrow(tmp_pilot)), ]
# write_xlsx(shuffled_tmp_pilot, "shuffled_pilot.xlsx")
# Now the order of data collection has been randomized

airplane_pilot <- read_excel("shuffled_pilot.xlsx",
                             col_types = c("numeric", "text", "text", "text"))
# The col "Distance" has been added into the excel file and loaded back to R as airplane_pilot

pilot_model <- lm(Distance ~ Nose*Middle*Rear, data=airplane_pilot)
output <- signif(summary(pilot_model)$coefficients, 4)
output[,] <- as.character(output[,])
kable(output)
```

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	166	11.6	14.31	1.551e-10
Nose1	7	16.41	0.4266	0.6754
Middle1	20.33	16.41	1.239	0.2332
Rear1	6.333	16.41	0.386	0.7046
Nose1:Middle1	-31.33	23.21	-1.35	0.1957
Nose1:Rear1	-16	23.21	-0.6895	0.5004
Middle1:Rear1	-28	23.21	-1.207	0.2451
Nose1:Middle1:Rear1	72.67	32.82	2.214	0.04169

```r
# Mean of beta1 through beta7
slope_mean <- round(mean(c(7, 20.3, 6.3, -31.3, -16, -28, 72.7)), 1)

# beta_se1 was calculated by shrinking the smallest beta estimate, intercept, by 33%
beta_mean <- c(168, rep(slope_mean, 7))
beta_se1 <- rep(7.8, 8)

replicates <- 2:10
power1 <- NA
for (i in 1:length(replicates)){
  power1[i] <- power_factorial_23(beta_mean,
                                  beta_se1,
                                  replicates[i])

}

beta_se2 <- rep(16.4, 8)

replicates <- 2:10
power2 <- NA
for (i in 1:length(replicates)){
```

```
  power2[i] <- power_factorial_23(beta_mean,
                                  beta_se2,
                                  replicates[i])
}

# beta_se3 was calculated by increasing the largest beta estimate by 25%
beta_se3 <- rep(41, 8)
replicates <- 2:10
power3 <- NA
for (i in 1:length(replicates)){
  power3[i] <- power_factorial_23(beta_mean,
                                  beta_se3,
                                  replicates[i])
}

all_power <- data.frame(power = c(power1, power2, power3),
                        beta_se = c(rep("7.8", length(power1)),
                        rep("16.4", length(power2)),
                        rep("41", length(power3))),
                        replicates = rep(replicates, 3))

ggplot(data=all_power, mapping = aes(x = replicates,
                                     y = power,
                                     color = beta_se,
                                     group = beta_se)) +
  geom_point() + geom_line()
```



Per sample size calculations, assuming a best case smaller standard error of 7.8 for β estimates, at least 6 replicates are required to get 80% power (1 replicate represents all 8 combinations of the factors). However, using the middle standard error estimate for β, 16.4, less than 50% power is obtained with 10 full replicates and barely 12% power is reached in 10 replicates in the case of a high β standard error of 41. These standard error values were calculated by reducing the lowest observed standard error by 33% and increasing the highest standard error value by 25%. The middle standard error value was kept the same as observed in the pilot study.

Because of feasibility, the most replicates that will be collected will be 12 (96 observations), which can still provide almost 100% power if there is a lower standard error of 7.8.

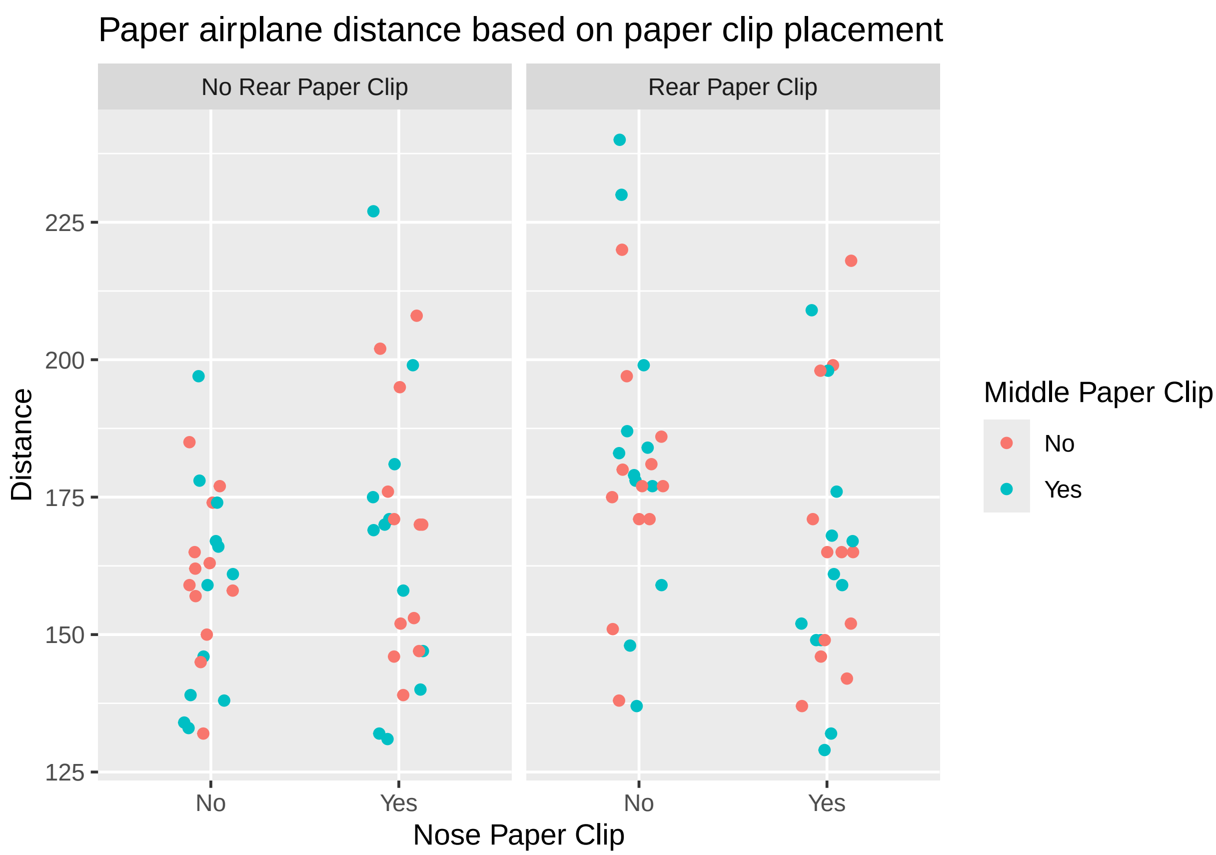
Since this data is intended for use as a pilot study, these calculations were done before collecting the real experimental data. Now, the required number of replicates to get sufficient power over a range of potential β standard error values is known.

Experimental data with 12 replicates:

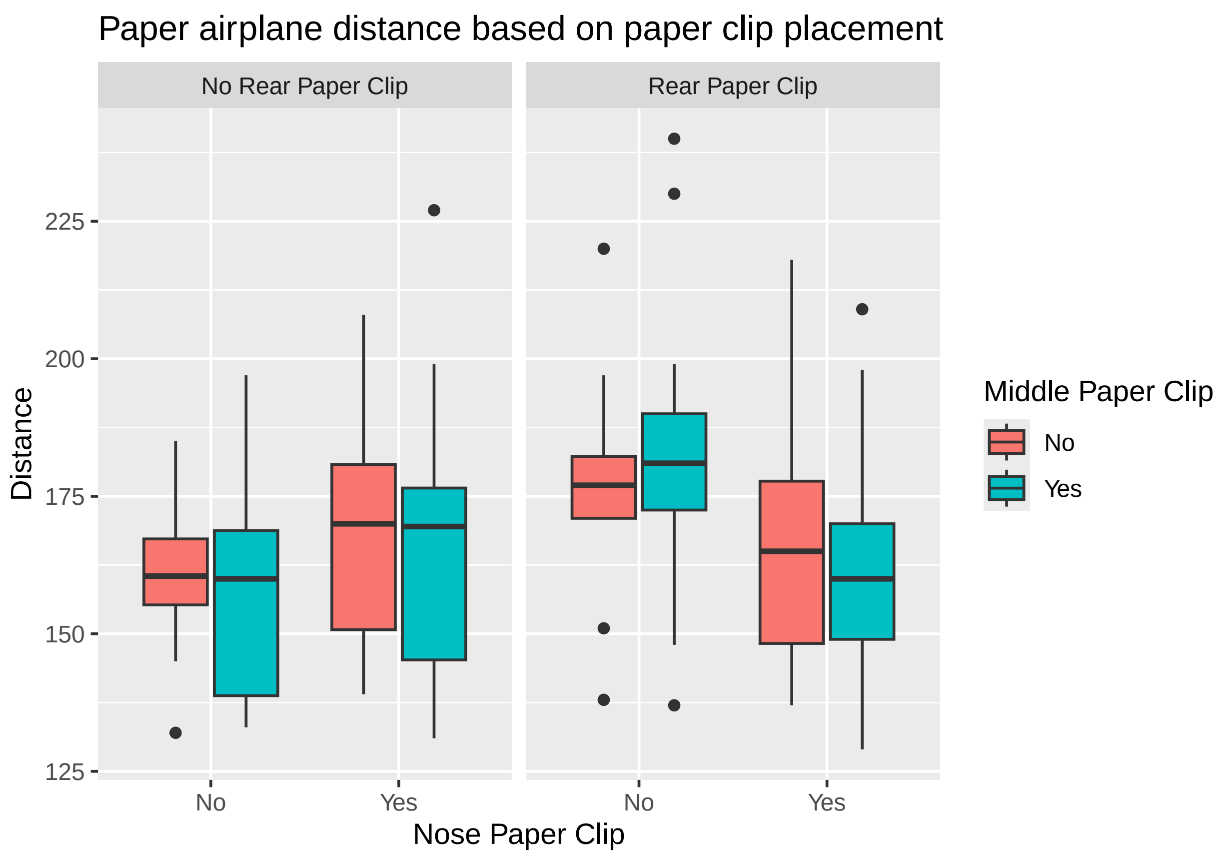
Visualizations:

```
# Randomization of the order of data collection:

# tmp_airplane <- read_excel("tmp_airplane.xlsx")
# shuffled_tmp_airplane <- tmp_airplane[sample(nrow(tmp_airplane)), ]
# write_xlsx(shuffled_tmp_airplane, "shuffled_airplane.xlsx")
# Now the order of data collection has been randomized

paper_airplane <- read_excel("shuffled_airplane.xlsx",
                             col_types = c("numeric", "text", "text", "text"))
```

```r
# The col "Distance" has been added into the excel file and loaded back to R as paper_airplane

ggplot(data=paper_airplane, mapping=aes(x=Nose, y=Distance, color=Middle)) +
  geom_jitter(width = 0.15, height = 0) +
  facet_grid(cols = vars(Rear),
             labeller = labeller(Rear = c("0" = "No Rear Paper Clip",
                                          "1" = "Rear Paper Clip"))) +
  ggtitle("Paper airplane distance based on paper clip placement") +
  scale_x_discrete(name = "Nose Paper Clip", labels = c("No", "Yes")) +
  scale_color_discrete(name = "Middle Paper Clip", labels = c("No", "Yes"))
```




```r
ggplot(data=paper_airplane, mapping=aes(x=Nose, y=Distance, fill=Middle)) +
  geom_boxplot() +
  facet_grid(cols = vars(Rear),
             labeller = labeller(Rear = c("0" = "No Rear Paper Clip",
                                          "1" = "Rear Paper Clip"))) +
  ggtitle("Paper airplane distance based on paper clip placement") +
  scale_x_discrete(name = "Nose Paper Clip", labels = c("No", "Yes")) +
  scale_fill_discrete(name = "Middle Paper Clip", labels = c("No", "Yes"))
```



Based on the visualizations of the data, there are not many apparent trends except for a slight increase in overall flight distance when the rear paper clip is present and the nose paper clip is not. It is more apparent in the boxplot that the presence of a middle paper clip creates more low flying distance values than other combinations.

Summary statistics:

```
library(dplyr)

summary_stats <- paper_airplane %>%
  group_by(Nose, Middle, Rear) %>%
  summarize(Mean = signif(mean(Distance), 4),
            Sd = signif(sd(Distance), 4),
            n = n())

kable(summary_stats)
```

Nose	Middle	Rear	Mean	Sd	n
0	0	0	160.6	14.36	12
0	0	1	177.0	20.51	12
0	1	0	157.7	20.08	12
0	1	1	183.4	29.73	12
1	0	0	169.1	22.96	12
1	0	1	167.2	25.46	12
1	1	0	166.7	28.03	12

Nose	Middle	Rear	Mean	Sd	n
1	1	1	162.4	23.74	12

From the summary statistics, it can be seen that the factor combination with the farthest average distance is when the middle and rear paper clips are attached to the airplane. However this combination also has the highest standard deviation, so these results had the most variation in distance out of all of the combinations. The combination with the shortest average distance was the group with no paper clips attached to the airplane at all. This group also had the smallest standard deviation, giving the most consistent results.

Analysis:

```
model1 <- lm(Distance ~ Nose*Middle*Rear, data=paper_airplane)
f_stat <- summary(model1)$fstatistic
pval <- pf(f_stat[1], df1=f_stat[2], df2=f_stat[3], lower.tail=FALSE)
```

Given that the overall model p-value is 0.1483, at significance level $\alpha = 0.05$ there is not statistically significance evidence to reject to the null hypothesis H_0 that $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = 0$, so the conclusion is that no combination of the factors (paper clip on the nose, middle, rear) has any impact on paper airplane flight distance.

```
output_table <- signif(summary(model1)$coefficients, 4)
output_table[,] <- as.character(output_table[,])

kable(output_table)
```

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	160.6	6.799	23.62	7.47e-40
Nose1	8.5	9.616	0.884	0.3791
Middle1	-2.917	9.616	-0.3033	0.7624
Rear1	16.42	9.616	1.707	0.0913
Nose1:Middle1	0.5	13.6	0.03677	0.9708
Nose1:Rear1	-18.25	13.6	-1.342	0.183
Middle1:Rear1	9.333	13.6	0.6863	0.4943
Nose1:Middle1:Rear1	-11.75	19.23	-0.611	0.5428

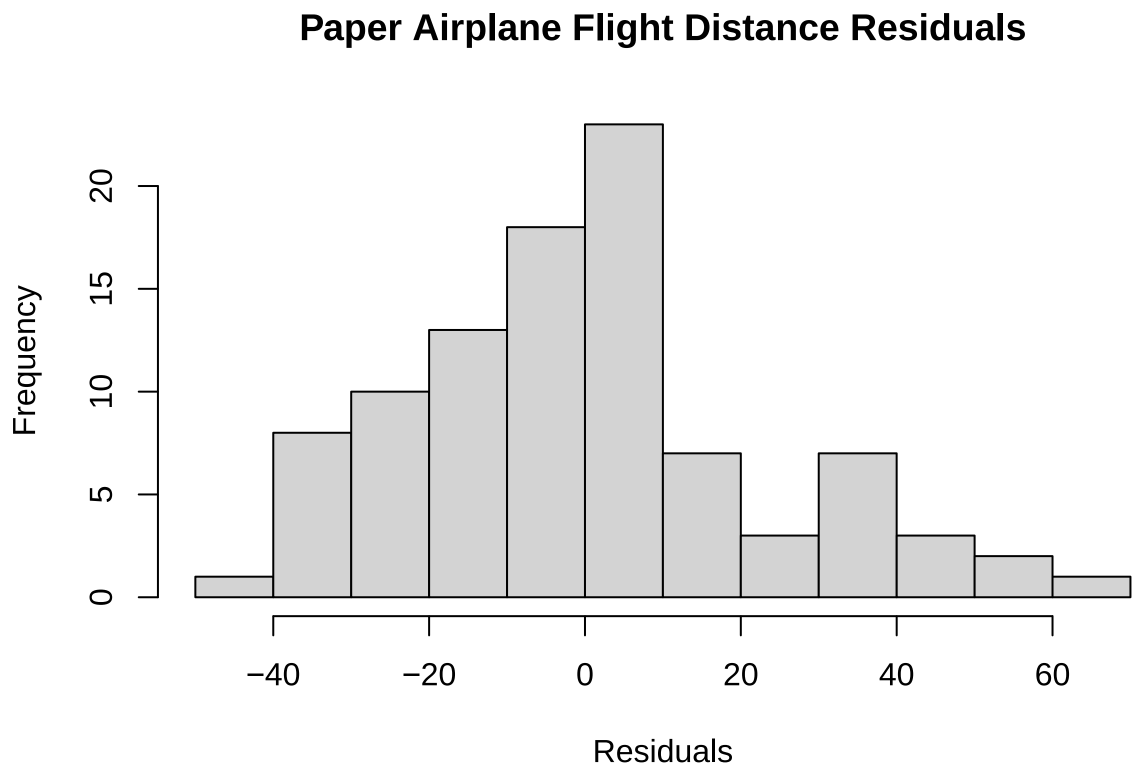
This conclusion remains constant with the individual coefficient p-values as well. At the Bonferroni-adjusted significance level of $\alpha_{Bonferroni} = \frac{0.05}{7} = 0.007$, the large p-values do not reject any of the null hypotheses, so none of the predictors have a statistically significant effect on flight distance.

However, the data runs the risk of not fulfilling the linear regression model's assumptions of normality, no structure to the data, and constant variance, so these should be checked for completeness.

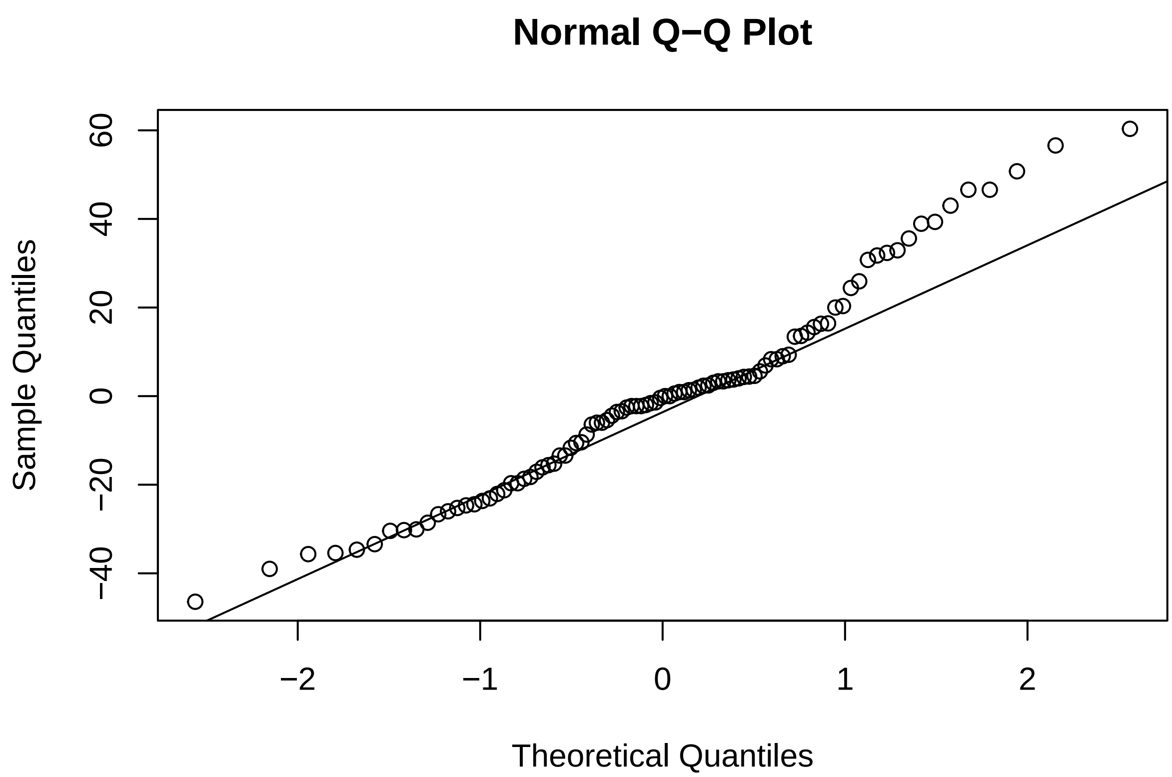
Check for normality:

Figure 1:

```
hist(model1$residuals, xlab="Residuals", main="Paper Airplane Flight Distance Residuals")
```




```r
qqnorm(model1$residuals)
qqline(model1$residuals)
```



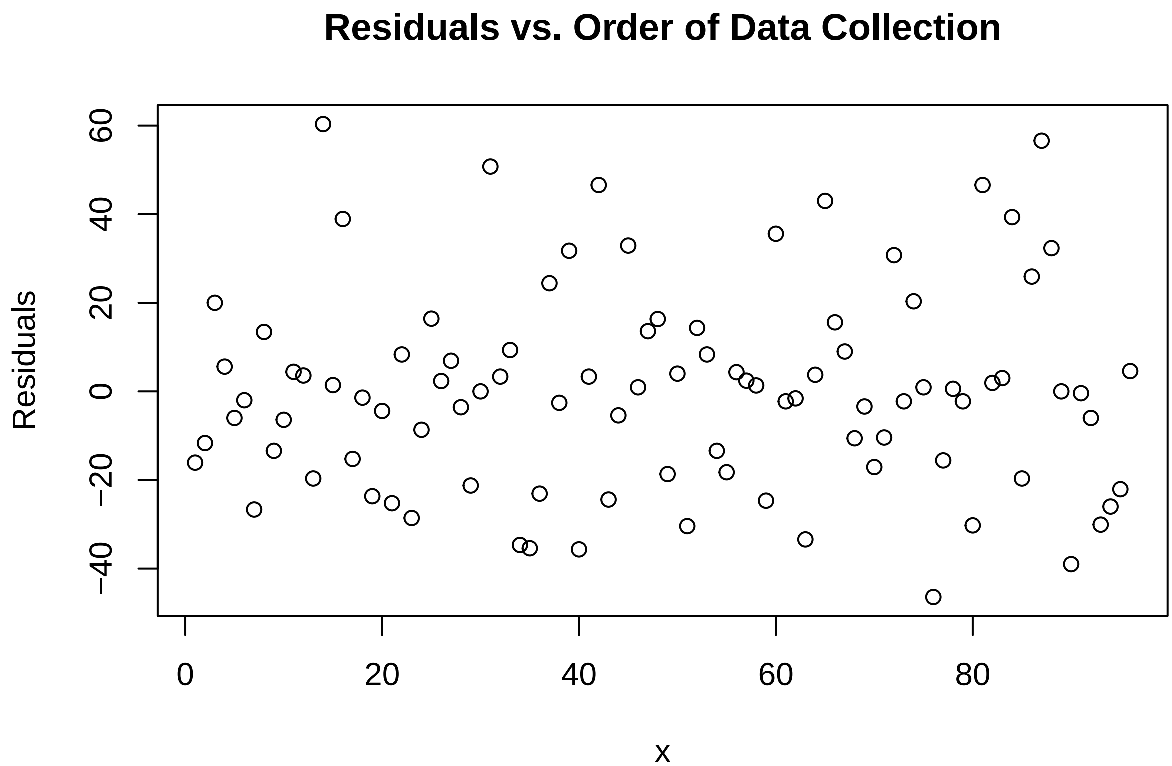

```r
shapiro_pval <- shapiro.test(model1$residuals)$p.value
```

The residuals of the data appear normally distributed in Figure 1, but to see this easier, the QQ-plot shows the residuals roughly following the theoretical normal values while they are negative, however they start to diverge from normality slightly as they become greater positive values. Finally, at $\alpha = 0.05$, the Shapiro-Wilk test concludes that there is statistically significant evidence that the residuals do not follow a normal distribution, rejecting H_0 with a p-value of 0.0208. Based on the analysis of the residuals, it is concluded that the paper airplane flight distance data is not normally distributed, thus violating linear regression's normality assumption. However, as will be seen later, this violation of normality may not be that big of an issue due to the central limit theorem.

Structure to the data:

Figure 2:

```r
x <- 1:length(model1$residuals)
plot(model1$residuals ~ x, ylab="Residuals", main="Residuals vs. Order of Data Collection")
```

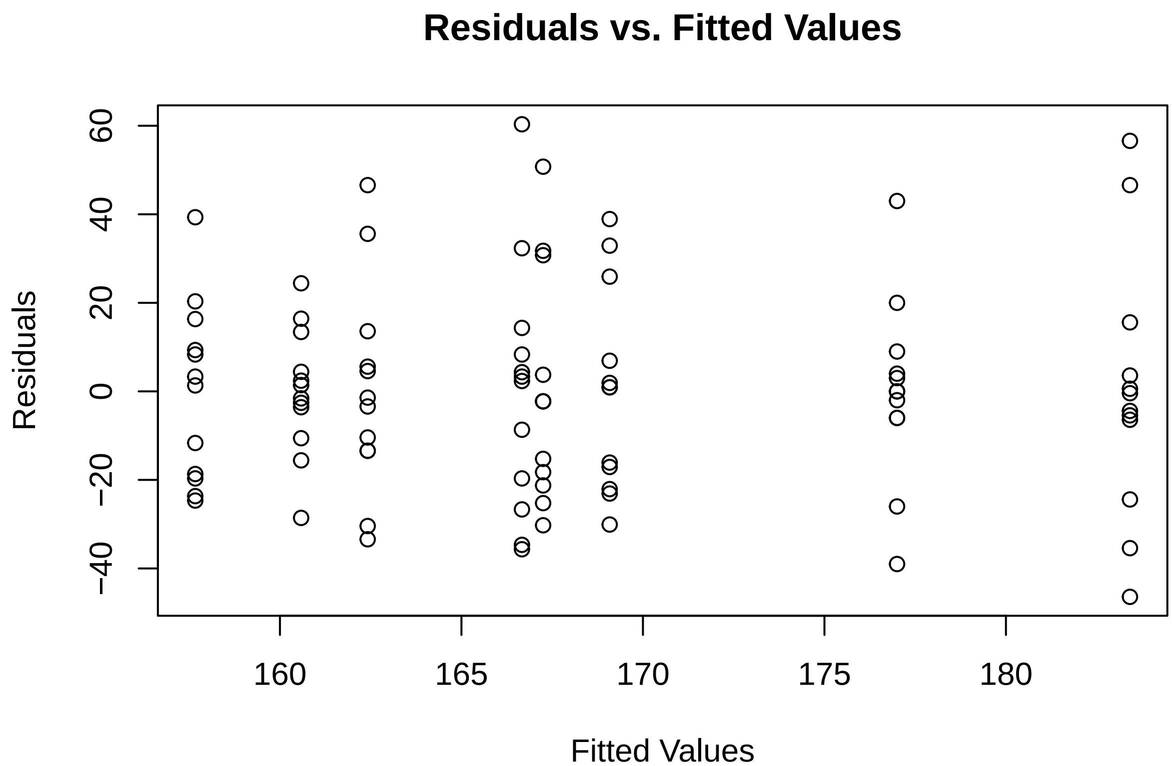


Based on Figure 2, there are no patterns, so it appears that there is no structure to the data.

Check for constant variance (homoscedasticity):

Figure 3:

```
plot(model1$residuals ~ model1$fitted.values,
     xlab="Fitted Values", ylab="Residuals", main="Residuals vs. Fitted Values")
```



Based on Figure 3, there is no concern about inequality of variance to the data.

Overall, the model assumptions were met except for normality, so a permutation test should be done to circumvent this assumption and maintain a normal Type I and Type II error rate.

Permutation test:

```
perm_f <- NA
reps <- 10000
for (i in 1:reps){
  perm_data <- paper_airplane
  perm_data$Distance <- sample(perm_data$Distance)
  perm_f[i] <- summary(lm(Distance ~ Nose*Middle*Rear, data=perm_data))$fstatistic[1]
}

perm_pval <- sum(perm_f >= f_stat[1]) / reps
```

From the permutation test, a p-value of 0.1494 is obtained, which is very similar to the p-value of 0.1483 from the standard test result. At $\alpha = 0.05$, H_0 is not rejected so the conclusion that the factors have no effect on paper airplane flight distance is upheld. Since the original p-value was not on the cusp of the rejection region, the permutation test outputs a similar test result. Additionally, with a large enough sample size like the one present in this study ($n = 96$), the the normality assumption is not super strict because of the central limit theorem, so the p-value of the permutation test is very similar to the original p-value. Therefore, no additional analysis of the individual effects is necessary (e.g. running multiple permutation tests) since it can be seen that none of the predictors are significant.

Discussion

This study assessed the impact of mass and weight distribution on paper airplane flight distance. By attaching paper clips to three spots on the paper airplane, I explored whether these differences in airplane architecture had any impact in how far the paper airplane was able to travel. The order of data collection was randomized through a sampling function, and a linear regression model and permutation test were employed for statistical analysis. The results suggest that weight distribution does not affect drag, lift, or subsequently flight distance of airplanes, but we know that this is quite the contrary in real aircraft. Although this study did not discover any statistically significant results, it is still a plausible method to to analyze how weight, drag, and lift factors affect the flight of real airplanes, and other studies have been more successful in their findings, discovering statistically significant evidence relating to airplane weight and nose construction [5].

A potential issue that could have altered my results was my throwing variability of the paper airplanes. Differences in the angle I threw the plane at, the strength I used to throw it, and consistency (or lack of) of my throws may have introduced variability into the experiment unrelated to the factors of interest. In the future, introducing blocks related to different throwers could help aid in reducing this unwanted variability. Furthermore, there was a breach in the assumption of normality needed for the linear regression model. After the data was collected, the structure of the data was analyzed and it was not normally distributed. However, given my sufficient sample size of $n = 96$, the violation of this assumption should not affect much of the experiment due to the central limit theorem, which claims that the regression coefficients approximate normality if the sample size is sufficiently large. Still, to circumvent the parametric assumption, a permutation test was done for completeness, however it ended in the same conclusion. The linear regression model's assumptions of constant variance and no structure to the data were upheld after looking at the residuals. Lastly, some limitations of the experiment were my inconsistent throwing technique, which gradually improved over time, and the deterioration of the paper airplane over time. The nose of the paper airplane became crumpled with successive trials, but the randomization of the order of data collection helped prevent a link of the wear and tear of the plane to a specific combination of factors. All in all, research points to the notion that weight of paper airplanes does affect flight distance, but more investigation will be required to find significant results within my specific experimental design.

References

[1] Sun Y, Kuprikov MY, Kuznetsova EL. Effect of flight range on the dimension of the main aircraft. INCAS Bulletin. 2020;12:201-9.

[2] Filippone A. Comprehensive analysis of transport aircraft flight performance. Progress in Aerospace Sciences. 2008 Apr 1;44(3):192-236.

[3] Harikumar A, Vibhute V. Aerodynamic Principles for Aircraft: A Study. Inte. J. Res. Appli. Sci. Eng. Technol. 2020;8:45-98.

[4] Pan J, Zhang Y. Calculation Method of the Ground Backflip Weight and Center of Gravity Limitation for Civil Aircraft. In2024 3rd International Symposium on Aerospace Engineering and Systems (ISAES) 2024 Mar 22 (pp. 256-259). IEEE.

[5] Puspita AN, Ambarwati L, Ludfiyanti E, Tejonugroho DP, Zen Y. Effect of paper weight, paper length, and nose of paper plane on aircraft mileage in paper airplane game.