Week 2 (LAB1): Paper Airplanes Design

PSTAT122: Design and Analysis of Experiments

Fall 2025

STUDENT NAME

- TA Daniel (NetID)

Due Date

Due Date: Sunday, October 12, 2025, 11:59 PM

1 Methods

- Start with a brief description of the aim of this lab
- Then give a description of how you are throwing the paper airplanes and measuring their distances flown
- Describe how and why you randomized the order of throws, and give an easy-to-read display of the order that you used.
- Describe any technical issues that may have come up, such as airplane deterioration, extreme wind, paperclip placement, etc., and how you dealt with them (even if you didn't have a great way to deal with any of it, just say so and describe any possible implications)

In this lab, we ran an experiment to test whether putting a paperclip on the nose of a paper airplane increases the distance it flies. We had 5 students each throw a paper airplane twice: once with a paperclip attached to the nose and once without. We measured the distance flown in feet for each throw. We designed a paired t-test study where one paper airplane was made for the entire group, and each student throws the same airplane once with and once without the paperclip. A tape measure was laid down, and the distance that the airplane flew was recorded in feet from the starting point.

We used the R function `sample` to randomize the order of the throws. This was done to minimize any potential bias that could arise from the order in which the throws were made, such as fatigue or learning effects. The randomized order of trials is shown below:

```r
trials <- c(
    "Person_1_with", "Person_1_without",
    "Person_2_with", "Person_2_without",
    "Person_3_with", "Person_3_without",
    "Person_4_with", "Person_4_without",
    "Person_5_with", "Person_5_without")

set.seed(123) # for reproducibility
ordered_trials <- sample(trials)
ordered_trials <- matrix(ordered_trials, ncol=1)
rownames(ordered_trials) <- 1:length(trials)
colnames(ordered_trials) <- c("Randomized Throw Order of Trials")
knitr::kable(ordered_trials, row.names = TRUE)
```

	Randomized Throw Order of Trials
1	Person_2_with
2	Person_5_without
3	Person_1_without
4	Person_4_without
5	Person_3_without
6	Person_5_with
7	Person_1_with
8	Person_4_with
9	Person_3_with
10	Person_2_without

2 Results

- Start by making a table in `R` of your raw data, with each person as a row, and three columns: with, without, and the difference. (**Hint**: you can use the `kable` function on your dataframe directly if you imported it via e.g. an Excel file; if you entered your data as two vectors, you can use the `cbind` function to put them together into a matrix, then use `kable` on that (but make sure to add appropriate rownames).
- At the bottom of this table or as a separate table, also show summary statistics for with, without AND the differences: `n`, `mean`, `sd`.

- Carefully write out your null and alternative hypotheses; if you use any mathematical symbols, carefully define what they represent
- Show your code for performing the hypothesis test. Present your output from the hypothesis test in a table.
- Write your results in text, with a conclusion in the context of the question at hand
- Include a sentence with the confidence interval, and briefly describe whether this is consistent with your hypothesis test.

```r
# Here is an example of how the raw data table should look (recorded in feet)
with <- c(30, 28, 32, 27, 29)
without <- c(25, 26, 24, 23, 22)
diff <- with - without

data <- cbind(with, without, diff)
rownames(data) <- c(
    "Person 1", "Person 2",
    "Person 3", "Person 4",
    "Person 5")

knitr::kable(data)
```

	with	without	diff
Person 1	30	25	5
Person 2	28	26	2
Person 3	32	24	8
Person 4	27	23	4
Person 5	29	22	7

```r
# Summary statistics
with_n <- length(with)
without_n <- length(without)
diff_n <- length(diff)

with_mean <- mean(with)
without_mean <- mean(without)
diff_mean <- mean(diff)

with_sd <- sd(with)
without_sd <- sd(without)
```

```r
diff_sd <- sd(diff)

summary_stats <-matrix(c(with_n, with_mean, with_sd,
                          without_n, without_mean, without_sd,
                          diff_n, diff_mean, diff_sd),
                        nrow=3, byrow=TRUE)
colnames(summary_stats) <- c("n", "mean", "sd")
rownames(summary_stats) <- c("With", "Without", "Difference")

knitr::kable(summary_stats)
```

	n	mean	sd
With	5	29.2	1.92353840617
Without	5	24.0	1.58113883008
Difference	5	5.2	2.38746727726

The mean of the differences is 5.2, meaning that on average, the paper airplanes flew further by 5.2 feet when the paperclip was attached. We now want to test whether this difference is statistically significant.

We set up our hypotheses as follows:

$$H_0 : \mu_d <= 0$$

$$H_a : \mu_d > 0$$

where μ_d is the true mean of the differences in distance flown with and without the paperclip.

```r
# Perform data analysis

diff <- with - without
d_bar <- mean(diff)
S <- sd(diff)
t0 <- d_bar/(S/sqrt(length(diff)))
CI_95lower <- d_bar - qt(0.975, df=length(diff)-1)*(S/sqrt(length(diff)))
CI_95upper <- d_bar + qt(0.975, df=length(diff)-1)*(S/sqrt(length(diff)))

t_table <- matrix(
    c(
        t0, length(diff)-1,
        pt(t0, df=length(diff)-1, lower.tail=FALSE),
        CI_95lower, CI_95upper),
```

```r
    nrow=1)
colnames(t_table) <- c(
    "t-statistic", "df", "p-value", "95% CI Lower", "95% CI Upper")
rownames(t_table) <- c("Paired t-test")
knitr::kable(t_table)
```

	t-statistic	df	p-value	95% CI Lower	95% CI Upper
Paired t-test	4.87024622022	4	0.004109114851	2.23556783493	8.16443216507

With a p-value of 0.0041091, at $\alpha = 0.05$, we reject the null hypothesis and conclude that there is significant evidence to suggest that attaching a paperclip to the nose of the paper airplane increases the distance it flies.

The 95% confidence interval for the mean difference in distance flown is (2.11, 8.29), which does not include 0, further supporting our conclusion.

Further note that this p-value was one-sided and the confidence interval was two-sided, which is why the p-value is less than 0.05 while the confidence interval does not include 0. A one-sided test is appropriate here because we are specifically interested in whether the paperclip increases the distance flown, not just whether it changes it in either direction. A two-sided hypothesis test would be appropriate if we were interested in any difference, whether an increase or decrease. Both are acceptable as long as the report is consistent with that throughout.

3 Discussion

- State any assumptions that are required for the statistical test that you performed. Briefly discuss whether you think the assumptions are reasonable (there is no need for formal checking yet, just your best guesses are fine).
- Describe any drawbacks of what you did, and of this study design in particular. Some of this may overlap with the "technical issues" you described in the Methods section, but regarding the study design, you could discuss issues of having just one paper airplane, the sample size (both in terms of the number of students and the number of times each student threw), placement of the paper clip, etc.

The assumptions required for our statistical test are that the differences follow a normal distribution (it is also correct to say that the distances themselves follow a normal distribution, but since we are doing a paired t-test, it is more relevant to consider the differences). Given our small sample size of 5, it is difficult to assess normality, but we can reasonably assume that the differences are approximately normally distributed based on the nature of the data. These throws also represent a random sample from all possible throws that could have been performed by these individuals, and the throws are independent of each other. Statements of independence among throwers is reasonable, as each person's throw does not influence another typically, but if you have a truly random sample, independence follow automatically from that.

Note that both of these assumptions are questionable given our small sample size, and we would ideally want to have a larger sample size to better assess normality and ensure the validity of our results. Plus, we did not randomly sample students to perform this experiment, so we cannot generalize our results to all students.

There are a number of possibilities to discuss here regarding drawbacks of the study design. First, we only used one paper airplane design, so our results may not generalize to other designs. Additionally, we only had 5 students participate, which is a small sample size and limits the power of our statistical test. Each student only threw the airplane once with and once without the paperclip, which introduces variability due to individual differences in throwing ability and technique. Having each student throw multiple times in each condition would help to average out this variability and provide a more accurate estimate of the true effect of the paperclip. The placement of the paperclip could also affect the results; if it was not placed consistently across all throws, this could introduce additional variability. Finally, external factors such as wind or fatigue could have influenced the results, and we did not control for these factors in our study design.