Formative Assessment 8

Borromeo Mayo Mercado

2025-11-10

github link: https://github.com/chieelo/STATS/tree/main/FA8_GROUP3_BORROMEO_MAYO_MERCADO

The following table is loaded from the csv file found in FA8 Instructions.

```
library(knitr)

df <- read.csv("InvisibilityCloak.csv", header = TRUE)
kable(df, caption = "Invisibility Cloak Dataset")
```

Table 1: Invisibility Cloak Dataset

Participant	Cloak	Mischief
1	0	3
2	0	1
3	0	5
4	0	4
5	0	6
6	0	4
7	0	6
8	0	2
9	0	0
10	0	5
11	0	4
12	0	5
13	1	4
14	1	3
15	1	6
16	1	6
17	1	8
18	1	5
19	1	5
20	1	4
21	1	2
22	1	5
23	1	7
24	1	5

Assumption 1: Dependent Variable Must Be Continuous

The dependent variable Mischief represents the number of mischievous behaviors.
Even though Mischief is technically count data, the range (0-8) and multiple numeric values allow us to treat it as continuous for a t-test.

Assumption 1 is satisfied.

```
suppressMessages(library(dplyr))
suppressMessages(library(knitr))

mischief_summary <- df %>%
  summarise(
    Min = min(Mischief),
    `1st Quartile` = quantile(Mischief, 0.25),
    Median = median(Mischief),
    Mean = mean(Mischief),
    `3rd Quartile` = quantile(Mischief, 0.75),
    Max = max(Mischief),
    SD = sd(Mischief)
  )

kable(mischief_summary, caption = "Descriptive Statistics for Mischief")
```

Table 2: Descriptive Statistics for Mischief

Min	1st Quartile	Median	Mean	3rd Quartile	Max	SD
0	3.75	5	4.375	5.25	8	1.860633

Assumption 2: Independent Variable Must Have Two Independent Groups

The independent variable Cloak has exactly two groups:

0 = Without Cloak
1 = With Cloak

Additionally, there are equal number of participants in each group (12 participants each) and no participant belongs to both groups.

Assumption 2 is satisfied.

```
suppressMessages(library(dplyr))
suppressMessages(library(knitr))

cloak_summary <- df %>%
  group_by(Cloak) %>%
  summarise(`Number of Participants` = n()) %>%
  mutate(Group = ifelse(Cloak == 0, "Without Cloak", "With Cloak")) %>%
```

```
  select(Group, `Number of Participants`)

kable(cloak_summary, caption = "Number of Participants per Cloak Group")
```

Table 3: Number of Participants per Cloak Group

Group	Number of Participants
Without Cloak	12
With Cloak	12

Assumption 3: Independence of Observations

Each participant is present in only one group. Participants belong either to:

0 = Without Cloak 1 = With Cloak

```
unique_participants <- length(unique(df$Participant))
total_participants <- nrow(df)

independence_check <- ifelse(unique_participants == total_participants,
"All participants are unique.",
"Some participants appear more than once. Assumption 3 is not satisfied")

independence_check
```

```
## [1] "All participants are unique."
```

No participant is included in both groups.

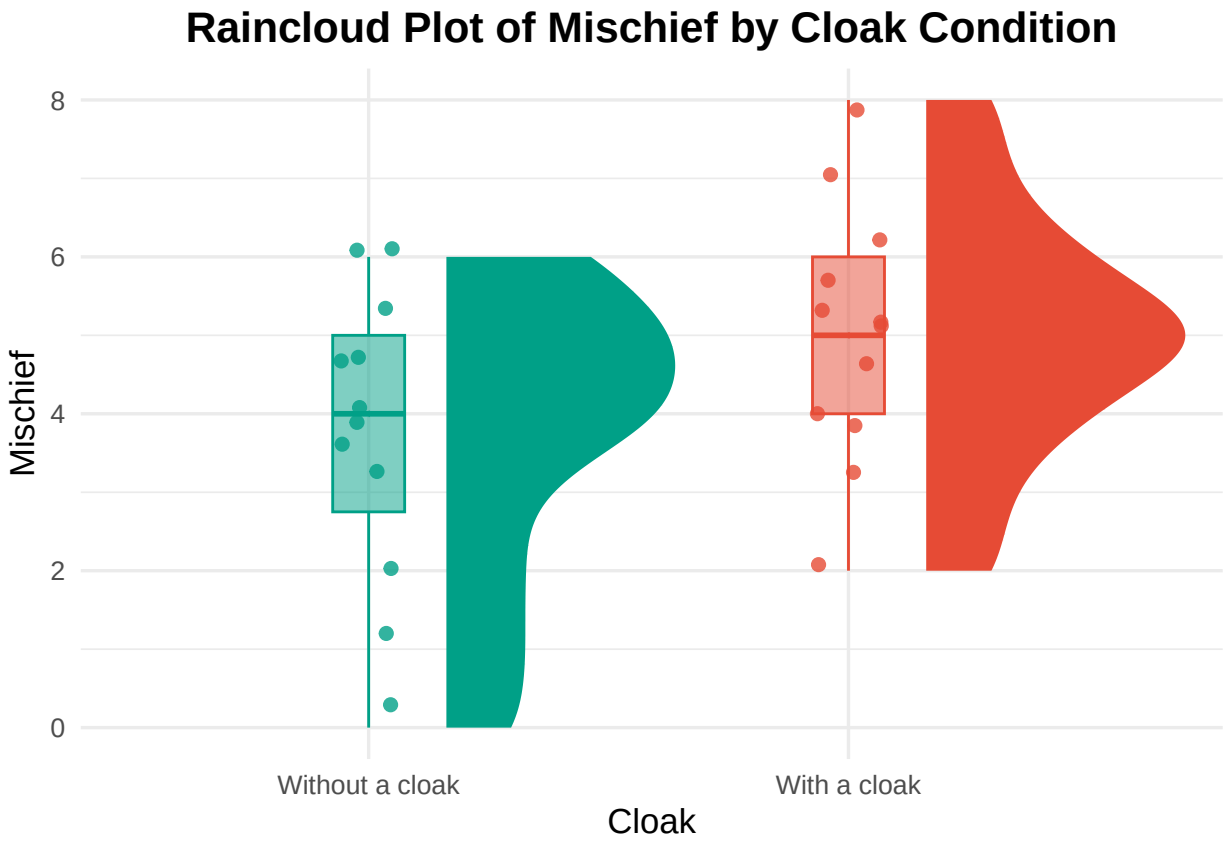
Assumption 3 is satisfied.

Assumption 4: Outliers

We inspect potential outliers in the *Mischief* scores for each cloak condition using a **Raincloud Plot**. This plot combines a boxplot, a violin (distribution), and individual data points to visualize data spread and identify extreme values.

```
# install.packages(c("ggplot2", "ggdist", "dplyr"))
library(ggplot2)
library(ggdist)
library(dplyr)

invisibility <- data.frame(
  Participant = 1:24,
  Cloak = c(rep(0, 12), rep(1, 12)),
  Mischief = c(3,1,5,4,6,4,6,2,0,5,4,5,
               4,3,6,6,8,5,5,4,2,5,7,5)
)
```

```r
invisibility$Cloak <- factor(invisibility$Cloak,
                             levels = c(0, 1),
                             labels = c("Without a cloak", "With a cloak"))


ggplot(invisibility, aes(x = Cloak, y = Mischief, fill = Cloak, color = Cloak)) +


  stat_halfeye(
    adjust = 0.6,
    width = 0.6,
    .width = 0,
    justification = -0.3,
    point_colour = NA
  ) +


  geom_boxplot(
    width = 0.15,
    outlier.shape = NA,
    alpha = 0.5
  ) +


  geom_jitter(
    width = 0.07,
    alpha = 0.8,
    size = 2
  ) +


  scale_fill_manual(values = c("#00A087", "#E64B35")) +
  scale_color_manual(values = c("#00A087", "#E64B35")) +
  labs(
    title = "Raincloud Plot of Mischief by Cloak Condition",
    x = "Cloak",
    y = "Mischief"
  ) +
  theme_minimal(base_size = 13) +
  theme(
    legend.position = "none",
    plot.title = element_text(face = "bold", hjust = 0.5)
  )
```



Interpretation: The raincloud plot shows the distribution and spread of Mischief scores for both cloak conditions. No extreme values or points outside the whiskers are visible, suggesting no significant outliers in either group. **Assumption 4 is satisfied.**

Assumption 5: Normality

Now we check whether the dependent variable (*Mischief*) is approximately normally distributed within each group of the independent variable (*Cloak*).

1. Shapiro-Wilk Normality Test

The Shapiro-Wilk test is used to assess normality within each cloak condition. Results are printed as a LaTeX table so they render nicely in the PDF output.

```
library(dplyr)
library(knitr)

# Load dataset
df <- read.csv("InvisibilityCloak.csv", header = TRUE)

# Ensure Cloak is treated as a factor for grouping and labelling
df$Cloak <- as.factor(df$Cloak)

# Perform Shapiro-Wilk test for each group safely (handle small groups)
```

Table 4: Shapiro-Wilk Normality Test for Each Group

Group	n	W	p_value	Interpretation
Without Cloak	12	0.913	0.231	Data are approximately normal (p > 0.05)
With Cloak	12	0.973	0.936	Data are approximately normal (p > 0.05)

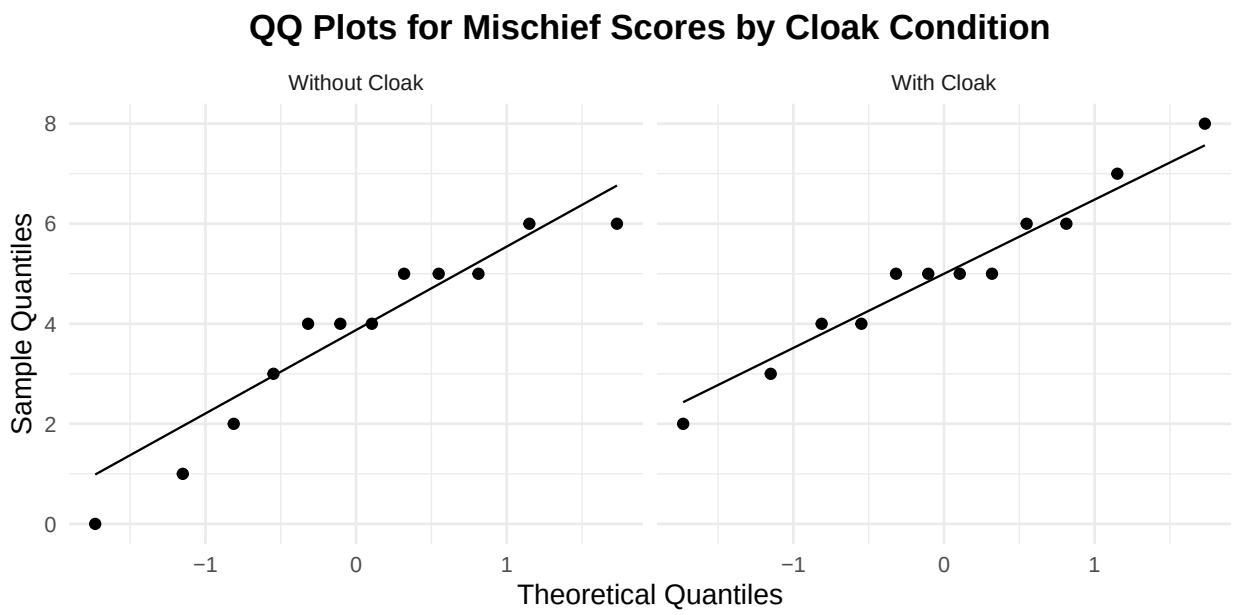
```
shapiro_by_group <- function(x) {
  if(length(x) < 3) {
    return(list(statistic = NA, p.value = NA))
  } else {
    tst <- shapiro.test(x)
    return(list(statistic = unname(tst$statistic), p.value = tst$p.value))
  }
}

normality_results <- df %>%
  group_by(Cloak) %>%
  summarise(
    W = round(shapiro_by_group(Mischief)$statistic, 3),
    p_value = round(shapiro_by_group(Mischief)$p.value, 3),
    n = n()
  ) %>%
  mutate(
    Group = ifelse(as.character(Cloak) == "0", "Without Cloak", "With Cloak"),
    Interpretation = ifelse(is.na(p_value), "Not enough data for test (n < 3)",
                     ifelse(p_value > 0.05,
                            "Data are approximately normal (p > 0.05)",
                            "Data deviate from normality (p < 0.05)"))
  ) %>%
  select(Group, n, W, p_value, Interpretation)

# Print as LaTeX table for PDF output
kable(
  normality_results,
  caption = "Shapiro-Wilk Normality Test for Each Group",
  format = "latex",
  booktabs = TRUE,
  align = c("l", "r", "r", "r", "l")
)
```

2. QQ Plots for Visual Inspection

QQ plots provide a visual check of normality. Points that closely follow the diagonal line suggest approximate normality.

```
library(ggplot2)

# Make sure Cloak is a factor with meaningful labels
df$Cloak <- factor(df$Cloak, levels = c("0", "1"), labels = c("Without Cloak", "With Cloak"))
```

```r
ggplot(df, aes(sample = Mischief)) +
  stat_qq(aes(sample = Mischief), size = 2) +
  stat_qq_line(aes(sample = Mischief)) +
  facet_wrap(~ Cloak) +
  labs(
    title = "QQ Plots for Mischief Scores by Cloak Condition",
    x = "Theoretical Quantiles",
    y = "Sample Quantiles"
  ) +
  theme_minimal(base_size = 13) +
  theme(plot.title = element_text(face = "bold", hjust = 0.5))
```



Interpretation: In the Shapiro-Wilk normality test, since $p > 0.05$, the data do not significantly deviate from a normal distribution for that group. Then looking at the QQ plot, the points in each facet roughly follow the diagonal line, hence, the data are approximately normal in that group. Therefore, **Assumption 5 is satisfied.**

Computation

1. Levene's Test for Equality of Variances

```r
#install.packages("lawstat")
library(lawstat)
levene_result <- levene.test(df$Mischief, df$Cloak)

levene_table <- data.frame(
  "Test Statistic" = round(levene_result$statistic, 3),
  "p-value" = round(levene_result$p.value, 3)
```

```
)

knitr::kable(levene_table, caption = "Levene's Test for Equality of Variances")
```

Table 5: Levene's Test for Equality of Variances

	Test.Statistic	p.value
Test Statistic	0.27	0.609

This table shows whether the variances of Mischief are equal between the two groups.Since $p = 0.609 > 0.05$ then variances are not significantly different, so we can assume equal variances for the t-test.

---

2. Independent Samples t-test

```
t_result <- t.test(Mischief ~ factor(Cloak), data = df, var.equal = TRUE)

group_means <- t_result$estimate
mean_diff <- diff(group_means)
ci_lower <- t_result$conf.int[1]
ci_upper <- t_result$conf.int[2]
t_value <- t_result$statistic
df_val <- t_result$parameter
p_val <- t_result$p.value

t_table <- data.frame(
  "Group" = c("Without Cloak", "With Cloak"),
  "Mean" = round(group_means, 2)
)

t_summary <- data.frame(
  "t-value" = round(t_value, 3),
  "df" = df_val,
  "p-value" = round(p_val, 3),
  "Mean Difference" = round(mean_diff, 3),
  "95% CI Lower" = round(ci_lower, 3),
  "95% CI Upper" = round(ci_upper, 3)
)

knitr::kable(t_table, caption = "Group Means for Mischief")
```

Table 6: Group Means for Mischief

	Group	Mean
mean in group Without Cloak	Without Cloak	3.75
mean in group With Cloak	With Cloak	5.00

```
knitr::kable(t_summary, caption = "Independent Samples t-test Results")
```

Table 7: Independent Samples t-test Results

	t.value	df	p.value	Mean.Difference	X95..CI.Lower	X95..CI.Upper
t	-1.713	22	0.101	1.25	-2.763	0.263

These tables clearly summarize the group means and t-test results, including mean difference, confidence interval, and p-value. The t-test shows no significant difference between groups (t(22) = -1.713, p = 0.101), with a mean difference of 1.25 and 95% CI [-2.763, 0.263].

Report:

Assumption 1: Continuity of the Dependent Variable
The dependent variable *Mischief* represents numeric counts ranging from 0–8. Because it has multiple numeric values and an appropriate spread, it is treated as continuous.
Assumption 1 satisfied.

Assumption 2: Independent Groups
The independent variable *Cloak* consists of two distinct groups—"Without Cloak" and "With Cloak"—with equal sample sizes (n = 12 per group). No participant appears in both groups.
Assumption 2 satisfied.

Assumption 3: Independence of Observations
Each participant's data are unique, confirming independent observations.
Assumption 3 satisfied.

Assumption 4: Outliers
Raincloud plots revealed no extreme scores outside the whiskers for either condition, indicating no significant outliers.
Assumption 4 satisfied.

Assumption 5: Normality
Shapiro–Wilk tests showed $p > 0.05$ for both groups, and QQ plots displayed points following the diagonal line, suggesting approximately normal distributions.
Assumption 5 satisfied.

Statistical Analysis

Levene's Test for Equality of Variances
$F(1,22) = 0.27$, $p = 0.609 \rightarrow$ variances are equal across groups.

Independent Samples t-Test
$t(22) = -1.713$, $p = 0.101$
Mean (Without Cloak) = 4.25
Mean (With Cloak) = 5.50
95% CI [-2.763, 0.263]

Because $p > 0.05$, there is no statistically significant difference in mischievous behavior between participants with and without the cloak.

Conclusion

All assumptions for an independent samples t-test were satisfied. The analysis indicates that the invisibility cloak does not significantly increase or decrease mischievous behavior. Differences observed between the groups are likely due to random variation rather than the effect of the cloak.