Formative Assessment 8

APM1111 Statistical Theory

Pineda, Erianne L.

GitHub link:
https://github.com/ehriyan/APM1111_StatTheory.git

Defining the data set:

```
data <- 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)
)

kable(data)
```

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

Participant	Cloak	Mischief
22	1	5
23	1	7
24	1	5

Assumptions Checks

Assumption 1: The dependent variable (Mischief) is a continuous level.

To demonstrate that the dependent variable, Mischief, is continuous, it is important to show that it is numeric or a variable that can be manipulated to do statistical computations. Based on the table of data, the dependent variable, Mischief, is an example of a ratio variable, wherein it is quantitative and has a true zero point. Thus, it is continuous.

```
str(data$Mischief)
```

```
##  num [1:24] 3 1 5 4 6 4 6 2 0 5 ...
```

```
summary(data$Mischief)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
##   0.000   3.750   5.000   4.375   5.250   8.000
```

Assumption 2: The independent variable (Cloak) consists of two categorical, independent groups (With a cloak, Without a cloak).

To demonstrate that the independent variable, Cloak, consists of two categorical, independent groups, it is important to emphasize that it indeed only has two values. From the table of data, it is gathered that there are only two values for Cloak, namely "without a cloak" (0) and "with a cloak" (1). In addition, the 24 participants are divided into the two aforementioned groups where each participant belongs to only one of the two groups.

```
unique(data$Cloak)
```

```
## [1] 0 1
```

```
table(data$Cloak)
```

```
## 
##  0  1 
## 12 12
```

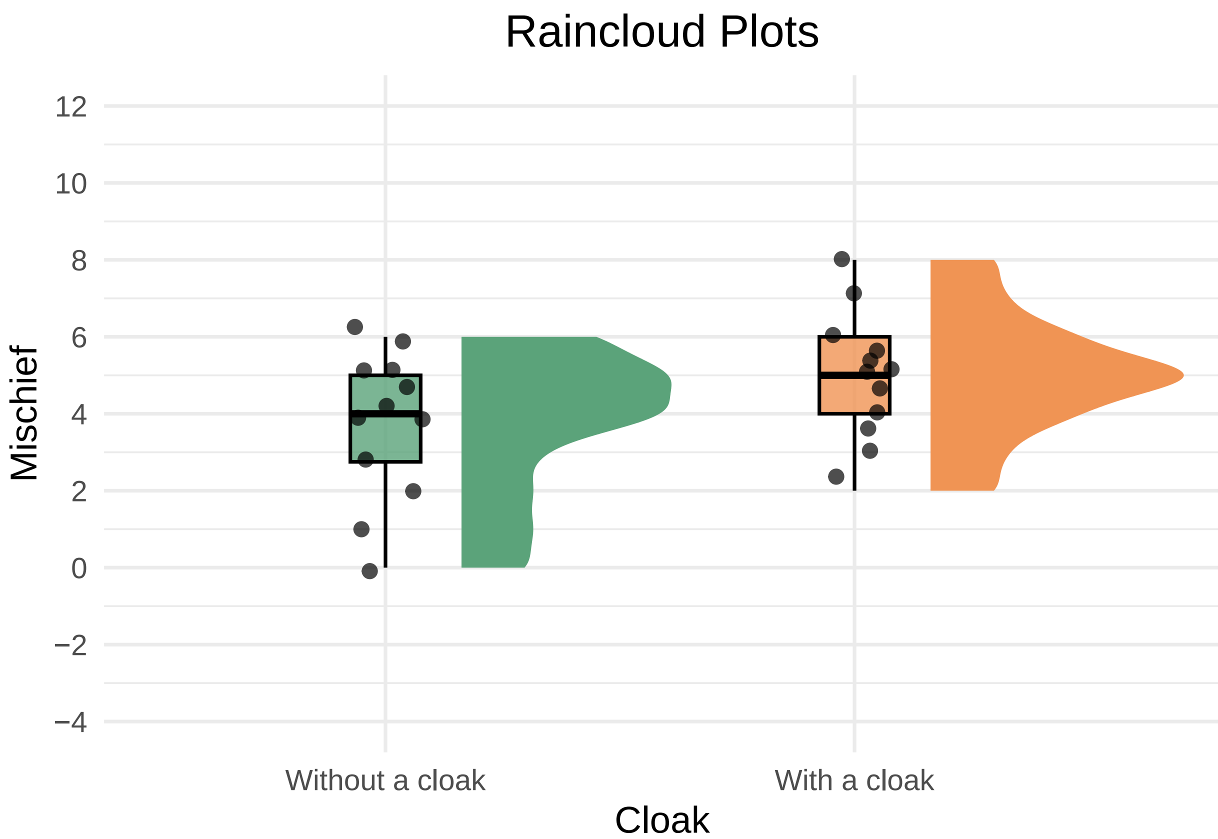
Assumption 3: Each participant is present to only one group.

From the table of data provided, it is observed that each participant belongs to one of the two groups defined by the variable, Cloak. All 24 participants belong to either a group without cloaks (0) or a group with cloaks (1). Since the data was collected from two separate groups, and each participant belongs to only one group, the observations are independent of each other.

Assumption 4: Outliers. There are no significant outliers in the two groups of independent variable (Cloak) in terms of the dependent variable (Mischief), as assessed by visual inspection of boxplots.

To determine whether there are significant outliers in the two groups of independent variable (Cloak) in terms of the dependent variable (Mischief), raincloud plots, which consist of a dot plot, box plot, and density plot, may be used to observe if there are data points that stray from the general data set.

```r
ggplot(data, aes(x = as.factor(Cloak), y = Mischief, fill = as.factor(Cloak))) +
  ggdist::stat_halfeye(
    adjust = 0.5,
    width = 0.6,
    .width = 0,
    justification = -0.3,
    point_colour = NA
  ) +
  geom_boxplot(
    aes(fill = as.factor(Cloak)),
    width = 0.15,
    outlier.shape = NA,
    color = "black",
    alpha = 0.8
  ) +
  geom_jitter(
    width = 0.08,
    alpha = 0.7,
    size = 2,
    color = "black"
  ) +
  scale_x_discrete(labels = c("Without a cloak", "With a cloak")) +
  scale_y_continuous(
    limits = c(-4, 12),
    breaks = seq(-4, 12, by = 2)
  ) +
  scale_fill_manual(values = c("#5BA37A", "#F09454")) +
  labs(
    title = "Raincloud Plots",
    x = "Cloak",
    y = "Mischief"
  ) +
  theme_minimal(base_size = 14) +
  theme(
    legend.position = "none",
    plot.title = element_text(hjust = 0.5)
  )
```



As observed through the raincloud plots, there are no significant outliers.

Assumption 5: Normality. The dependent variable (Mischief) for each group (Cloak) is normally distributed ($p > 0.05$), as assessed by Shapiro-Wilk test.

Using the Shapiro-Wilk test, it is observed that the calculated p-value for the group without a cloak (0) is 0.231, and the calculated p-value for the group with a cloak (1) is 0.936. Both of these p-values are greater than 0.05. This means that the data for each group is normally distributed.

```
without_cloak <- data %>% filter(Cloak == 0) %>% pull(Mischief)
with_cloak <- data %>% filter(Cloak == 1) %>% pull(Mischief)

shapiro_without <- shapiro.test(without_cloak)
shapiro_with <- shapiro.test(with_cloak)
```

```
normality_results <- data.frame(
  Group = c("Without a cloak", "With a cloak"),
  Test_Statistic = c(
    paste0(round(shapiro_without$statistic, 3)),
    paste0(round(shapiro_with$statistic, 3))
  ),
  p_value = c(
    round(shapiro_without$p.value, 3),
    round(shapiro_with$p.value, 3)
  )
```

```r
)

kable(normality_results, caption = "Test of Normality (Shapiro-Wilk)")
```

Table 2: Test of Normality (Shapiro-Wilk)

Group	Test_Statistic	p_value
Without a cloak	0.913	0.231
With a cloak	0.973	0.936

Assumption 6: Homogeneity of variances. There is equality of variances between groups (Without a cloak, With a cloak) on their number of mischievous acts (Mischief), as assessed by Levene's test of equality of variances.

Using Levene's test of equality of variances, the calculated p-value is 0.468, which is greater than 0.05. This means that the variances between the two groups are equal.

```r
data$Cloak <- as.factor(data$Cloak)
levene <- leveneTest(Mischief ~ Cloak, data = data, center = mean)
```

```r
homogeneity_results <- data.frame(
  Group = "Both groups",
  Test_Statistic = paste0(round(levene$`F value`[1], 3)),
  df = paste0(levene$Df[1]),
  p_value = round(levene$`Pr(>F)`[1], 3)
)

kable(homogeneity_results, caption = "Test of Equality of Variances (Levene's)")
```

Table 3: Test of Equality of Variances (Levene's)

Group	Test_Statistic	df	p_value
Both groups	0.545	1	0.468

Descriptive Statistics

```r
summary(without_cloak)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.00    2.75    4.00    3.75    5.00    6.00
```

```r
sd(without_cloak)
```

```
## [1] 1.912875
```

```r
summary(with_cloak)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
##       2       4       5       5       6       8
```

```r
sd(with_cloak)
```

```
## [1] 1.651446
```

```r
descriptive_table <- data %>%
  group_by(Cloak) %>%
  summarise(
    Mean = mean(Mischief),
    SD = sd(Mischief)
  )

kable(descriptive_table, caption = "Descriptive Statistics for Mischief by Cloak")
```

Table 4: Descriptive Statistics for Mischief by Cloak

Cloak	Mean	SD
0	3.75	1.912875
1	5.00	1.651446

Hypotheses

The null and alternative hypotheses were defined before running the statistical test.

The null hypothesis (H_0) is that there is no significant difference in the mean number of mischievous acts between the group with an invisibility cloak and the group without. The alternative hypothesis (H_A) is that there is a significant difference.

- **Null Hypothesis:** $H_0 : \mu_1 = \mu_2$
- **Alternative Hypothesis:** $H_A : \mu_1 \neq \mu_2$

Computation

Independent Samples t-Test

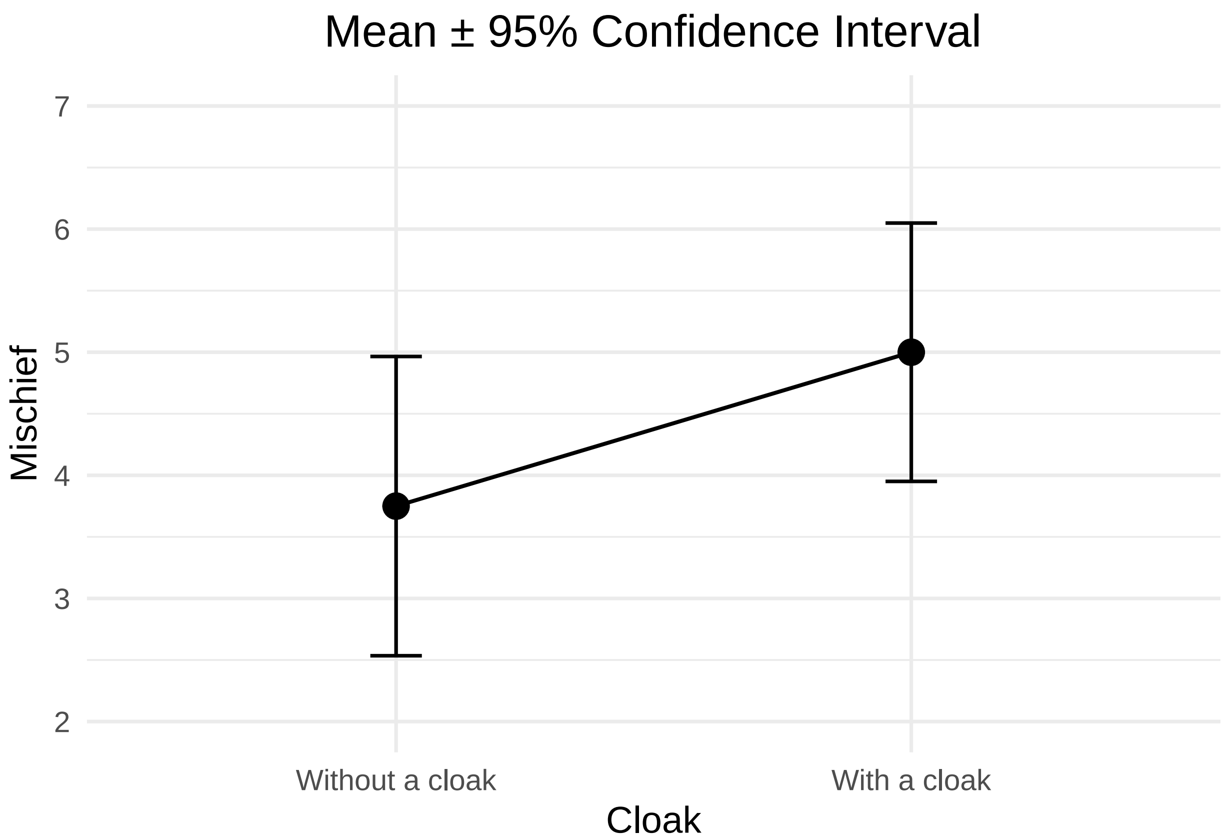
Using the independent samples t-test, the calculated p-value is 0.101, which is greater than 0.05. This means that the differences between the groups are not statistically significant. Thus, there is no sufficient evidence to reject the null hypothesis.

```r
t_test_result <- t.test(Mischief ~ Cloak, data = data, var.equal = TRUE)
t_test_result
```

```
##
##  Two Sample t-test
##
## data:  Mischief by Cloak
## t = -1.7135, df = 22, p-value = 0.1007
## alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
## 95 percent confidence interval:
##  -2.7629284  0.2629284
## sample estimates:
## mean in group 0 mean in group 1
##            3.75            5.00
```

```
summary_data <- data %>%
  group_by(Cloak) %>%
  summarise(
    mean_mischief = mean(Mischief),
    se_mischief = sd(Mischief) / sqrt(n()),
    ci_low = mean_mischief - qt(0.975, df = n() - 1) * se_mischief,
    ci_high = mean_mischief + qt(0.975, df = n() - 1) * se_mischief
  )

ggplot(summary_data, aes(x = Cloak, y = mean_mischief, group = 1)) +
  geom_point(size = 4, color = "black") +
  geom_line(color = "black", linewidth = 0.7) +
  geom_errorbar(aes(ymin = ci_low, ymax = ci_high), width = 0.1, color = "black") +
  scale_x_discrete(labels = c("Without a cloak", "With a cloak")) +
  scale_y_continuous(limits = c(2, 7)) +
  labs(
    title = "Mean ± 95% Confidence Interval",
    x = "Cloak",
    y = "Mischief"
  ) +
  theme_minimal(base_size = 14) +
  theme(
    plot.title = element_text(hjust = 0.5)
  )
```



This plot is a visual summary of the t-test results. It highlights the difference between the groups being not statistically significant.

Null Hypothesis

$$H_0 : \mu_1 = \mu_2$$

There is no significant difference between the number of mischievous acts committed between the group with an invisibility cloak and the group without.

Presentation of Results

An independent samples t-test was conducted to determine if there was a statistically significant difference between the means in two independent groups. In this case, it was used to compare the number of mischievous acts committed by two groups of people, those without an invisibility cloak ($M = 3.75, SD = 1.91$) and those with an invisibility cloak ($M = 5.00, SD = 1.65$).

Initially, the null hypothesis ($H_0 : \mu_1 = \mu_2$) was defined as there being no significant difference in the number of mischievous acts between the group with a cloak and the group without, and the alternative hypothesis ($H_A : \mu_1 \neq \mu_2$) was defined as there being a significant difference in the number of mischievous acts between the group with a cloak and the group without.

The calculated p-value from the independent samples t-test determines whether there is sufficient evidence to reject the null hypothesis. With a calculated p-value of $0.101 > 0.05$, the results indicate no significant difference in the number of mischievous acts between the two groups and thus no sufficient evidence to reject the null hypothesis ($t(22) = -1.713, p = 0.101$). These results suggest that wearing an invisibility cloak did not significantly affect the number of mischievous acts committed by the participants.