Module 4 | Practice 4

Trang Tran

CPS, Northeastern University

ALY6010 | Probability Theory and Introductory Statistics

Patrick McQuillan

Mar 18, 2023

Part I: "cats" dataset

```
> skim(cats)
── Data Summary ────────────────────────
                           Values
Name                       cats
Number of rows             144
Number of columns          3
_______________________
Column type frequency:
  factor                   1
  numeric                  2
_______________________
Group variables            None

── Variable type: factor ──────────────────────────────────────────────────────
  skim_variable n_missing complete_rate ordered n_unique top_counts
1 Sex                   0             1 FALSE          2 M: 97, F: 47

── Variable type: numeric ─────────────────────────────────────────────────────
  skim_variable n_missing complete_rate  mean    sd  p0  p25  p50   p75 p100 hist
1 Bwt                   0             1  2.72 0.485 2   2.3   2.7  3.02  3.9 ▃▇▃▁▁
2 Hwt                   0             1 10.6  2.43  6.3 8.95 10.1 12.1  20.5 ▅▇▂▁▁
```

```
> head(cats)
  Sex Bwt Hwt
1   F 2.0 7.0
2   F 2.0 7.4
3   F 2.0 9.5
4   F 2.1 7.2
5   F 2.1 7.3
6   F 2.1 7.6
```

Figure 1: Data Summary

Question: do male and female cat samples have the same body weight ("Bwt")?

We take a two-sample two-tailed t-test for the mean body weight of the male and female cats after filtering two separate groups of male and female.

```
        Welch Two Sample t-test

data:  male$Bwt and female$Bwt
t = 8.7095, df = 136.84, p-value = 0.000000000000008831
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.4177242 0.6631268
sample estimates:
mean of x mean of y
 2.900000  2.359574
```

H_0: Means' difference = 0

H_1: Means' difference ≠ 0

CI: 95%, α = 0.05

As we can see, this two-tailed t-test yields a p-value of much less than the significance level (α) of 0.05. Thus, we firmly can reject the null hypothesis and conclude that there is a significant difference between the average body weight of male and female cats. The sample means of male and female cats are ~2.9 and ~2.36 respectively. The 95% confidence interval for the difference of two means is from 0.4177 to 0.6631.

Part 2: Whether meditation has an effect on sleep quality?

We proceed with a paired t-test on a sample of 10 students and their average sleeping quality scores in the week before the meditation workshop and the week following the meditation workshop.

- Null hypothesis: the mean sleep quality scores before and after the mediation workshop are the same. ($\mu_D = 0$)
- Alternative hypothesis (claim): the mean sleep quality scores after the workshop are higher than before the workshop. ($\mu_D < 0$)

```
        Paired t-test

data:  before$bef and after$aft
t = -1.9481, df = 9, p-value = 0.04161
alternative hypothesis: true mean difference is less than 0
95 percent confidence interval:
        -Inf -0.03659503
sample estimates:
mean difference
          -0.62
```

CI: 95%, α ***= 0.05***

This left-tailed paired t-test yields a p-value of 0.04161, which is less than the significance level (α) of 0.05. Hence, we reject the null hypothesis and conclude that the mean sleep quality scores after the meditation workshop are significantly higher than before the workshop. It supports the claim that meditation has a good effect on sleep quality with a 95% of confidence.

If we change the level of significance to 0.1 which is equivalent to a 90% of confidence level, the p-value is still less than 0.1, so we would still reject the null hypothesis.

We chose to use the dependent samples t-test because we are comparing two groups of data that are dependent on each other. Each student's sleep quality before and after the workshop is dependent on the others. The t-test assumes that the data are normally distributed and have equal variances, which we can assume based on the sample size and the similarity of the two groups variances. Additionally, we chose the one-tailed test because we were interested in whether the sleep quality scores after the workshop were higher than before.